

10/040438
01/09/02

Date: 01/09/2002
Enclosures
92637-24/NWH/mis/January 8, 2002

PATENT OFFICE
JAPANESE GOVERNMENT



This is to certify that the annexed is a true copy of the
following application as filed with this Office.

DATE OF APPLICATION : January 10, 2001

APPLICATION NUMBER : Patent Application 2001-002070

APPLICANT(S) : MONOLITH CO., LTD.

November 26, 2001

Commissioner, Patent Office

Kozo OIKAWA

日 本 国 特 許 庁
JAPAN PATENT OFFICE

-1C826 U.S. PTO

10/040438

01/09/02

出証番号 出証特200.1-3102676

【書類名】 特許願

【整理番号】 MN-0124

【提出日】 平成13年 1月10日

【あて先】 特許庁長官殿

【国際特許分類】 H04B 14/00

【発明者】

 【住所又は居所】 東京都江戸川区西葛西 5 - 1 0 - 2 6

 【氏名】 品川 嘉久

【発明者】

 【住所又は居所】 千葉県千葉市稲毛区天台 4 - 2 - 1 2

 【氏名】 永嶋 浩樹

【発明者】

 【住所又は居所】 東京都港区麻布十番 1 - 7 - 3 株式会社モノリス内

 【氏名】 秋吉 仰三

【発明者】

 【住所又は居所】 東京都港区麻布十番 1 - 7 - 3 株式会社モノリス内

 【氏名】 秋吉 信雄

【特許出願人】

 【識別番号】 396001980

 【氏名又は名称】 株式会社モノリス

【代理人】

 【識別番号】 100105924

 【弁理士】

 【氏名又は名称】 森下 賢樹

 【電話番号】 0422-23-7415

【先の出願に基づく優先権主張】

 【出願番号】 特願2000-360640

 【出願日】 平成12年11月28日

【手数料の表示】

【予納台帳番号】 054977

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9717729

【プルーフの要否】 要

【書類名】 明細書

【発明の名称】 映像効果方法と装置および画像補間方法と装置

【特許請求の範囲】

【請求項 1】 第 1 画像上に設けられたメッシュの格子点に対応する第 2 画像上の点を画像マッチングによって検出する工程と、

その結果をもとに、第 1 画像上のメッシュを構成する起点ポリゴンに、第 2 画像上にて対応する終点ポリゴンを定義する工程と、

を含むことを特徴とする映像効果方法。

【請求項 2】 前記検出する工程は、第 1 画像と第 2 画像の間で画素単位のマッチング計算を行う請求項 1 に記載の方法。

【請求項 3】 前記検出する工程は、第 1 画像について二次元的な探索を行って検出した特異点と、第 2 画像について二次元的な探索を行って検出した特異点との対応をもとに画素単位のマッチング計算を行う請求項 2 に記載の方法。

【請求項 4】 前記検出する工程は、第 1 画像と第 2 画像を、それぞれ前記特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得する請求項 3 に記載の方法。

【請求項 5】 前記定義する工程は、起点ポリゴンと終点ポリゴンの対応関係を対応点ファイルに出力し、このファイルを第 1 画像および第 2 画像と関連づけて保持する請求項 1 から 4 のいずれかに記載の方法。

【請求項 6】 前記対応点ファイルを利用し、第 1 画像と第 2 画像の中間画像をポリゴン単位の補間計算にて生成する工程をさらに含む請求項 5 に記載の方法。

【請求項 7】 前記生成する工程は、起点ポリゴン内部の点に対応する終点ポリゴン内部の点を導出する請求項 6 に記載の方法。

【請求項 8】 起点ポリゴンを三角形に分割し、各三角形に関するベクトルの内分にて終点ポリゴン内部の点を求める請求項 7 に記載の方法。

【請求項 9】 第 1 画像および第 2 画像を取得する画像入力部と、

第1画像と第2画像のマッチングを計算するマッチングプロセッサとを含み、このマッチングプロセッサは、第1画像上のメッシュを構成する起点ポリゴンに、第2画像上にて対応する終点ポリゴンを定義することにより、対応点ファイルを生成することを特徴とする映像効果装置。

【請求項10】 前記マッチングプロセッサは、第1画像について二次元的な探索を行って検出した特異点と、第2画像について二次元的な探索を行って検出した特異点との対応をもとに画素単位のマッチング計算を行う請求項9に記載の装置。

【請求項11】 前記マッチングプロセッサは、第1画像と第2画像を、それぞれ前記特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得する請求項10に記載の装置。

【請求項12】 前記対応点ファイルを外部へ送出するための通信部をさらに含む請求項9から11のいずれかに記載の装置。

【請求項13】 第1画像、第2画像およびそれらの間の対応点ファイルを取得し、それらふたつの画像の中間画像を生成する中間画像生成部を含み、前記対応点ファイルが第1画像上にとられたメッシュの格子点と、それに対応する第2画像上の点の位置関係を記述していることを特徴とする映像効果装置。

【請求項14】 前記中間画像、前記第1画像、前記第2画像のタイミングを調整してそれらを表示する表示部をさらに含む請求項13に記載の装置。

【請求項15】 第1画像上に設けられたメッシュの格子点に対応する第2画像上の点を画像マッチングによって検出する工程と、

その結果をもとに、第1画像上のメッシュを構成する起点ポリゴンに、第2画像上にて対応する終点ポリゴンを定義する工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【請求項16】 第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第 1 画像と第 2 画像の中間画像を生成する工程と、
その中間画像の生成に当たり、電子商取引における商品プレゼンテーション機能を提供する工程と、
を含むことを特徴とする画像補間方法。

【請求項 1 7】 第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第 1 画像と第 2 画像の中間画像を生成する工程と、
その中間画像の生成に当たり、映像効果に関する機能を提供する工程と、
を含むことを特徴とする画像補間方法。

【請求項 1 8】 第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第 1 画像と第 2 画像の中間画像を生成する工程と、
その中間画像の生成に当たり、動画像の生成に関するユーティリティ機能を提供する工程と、
を含むことを特徴とする画像補間方法。

【請求項 1 9】 第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第 1 画像と第 2 画像の中間画像を生成する工程と、
その中間画像の生成を、モバイル機器への適用を前提とした機能のもとで行う工程と、

を含むことを特徴とする画像補間方法。

【請求項 2 0】 前記商品プレゼンテーション機能は、ユーザからの指示にしたがって商品を回転して表示する機能を含む請求項 1 6 に記載の方法。

【請求項 2 1】 前記商品プレゼンテーション機能は、商品画像を二方向に補間して得られた中間画像を表示する機能を含む請求項 1 6 に記載の方法。

【請求項 2 2】 前記映像効果に関する機能は、3 以上の画像の中から順次

ふたつの画像を第 1 画像および第 2 画像として選定し中間画像を連続的に表示する機能を含む請求項 1 7 に記載の方法。

【請求項 2 3】 前記映像効果に関する機能は、処理の対象となる画像を二方向に補間して得られた中間画像を表示する機能を含む請求項 1 7 に記載の方法。

【請求項 2 4】 前記ユーティリティ機能は、生成すべき中間画像の枚数を指定する機能を含む請求項 1 8 に記載の方法。

【請求項 2 5】 前記生成すべき中間画像の枚数に関する指定は G U I を介してユーザから取得される請求項 2 4 に記載の方法。

【請求項 2 6】 前記モバイル機器への適用を前提とした機能は、第 1 画像または第 2 画像の一方のみと前記対応情報をもとに生成された中間画像を表示する機能である請求項 1 9 に記載の方法。

【請求項 2 7】 第 1 画像または第 2 画像の一方のみと前記対応情報をストリーミング処理して前記中間画像が生成される請求項 2 6 に記載の方法。

【請求項 2 8】 画像を補間する装置であって、

第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得し、その対応情報もとに第 1 画像と第 2 画像の中間画像を生成する中間画像生成部を含み、

当該中間画像生成部は、中間画像の生成に当たり、電子商取引における商品プレゼンテーション機能を提供することを特徴とする画像補間装置。

【請求項 2 9】 画像を補間する装置であって、

第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得し、その対応情報もとに第 1 画像と第 2 画像の中間画像を生成する中間画像生成部を含み、

当該中間画像生成部は、中間画像の生成に当たり、映像効果に関する機能を提供することを特徴とする画像補間装置。

【請求項 3 0】 画像を補間する装置であって、

第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得し、その対応情報も

とに第 1 画像と第 2 画像の中間画像を生成する中間画像生成部を含み、

当該中間画像生成部は、中間画像の生成に当たり、動画像の生成に関するユーティリティ機能を提供することを特徴とする画像補間装置。

【請求項 3 1】 画像を補間する装置であって、

第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得し、その対応情報もとに第 1 画像と第 2 画像の中間画像を生成する中間画像生成部を含み、

当該中間画像生成部は、中間画像の表示を、モバイル機器への適用を前提とした機能のもとで行うことを特徴とする画像補間装置。

【請求項 3 2】 前記商品プレゼンテーション機能は、ユーザからの指示にしたがって商品を回転して表示する機能を含む請求項 2 8 に記載の装置。

【請求項 3 3】 前記商品プレゼンテーション機能は、商品画像を二方向に補間して得られた中間画像を表示する機能を含む請求項 2 8 に記載の装置。

【請求項 3 4】 前記映像効果に関する機能は、3 以上の画像の中から順次ふたつの画像を第 1 画像および第 2 画像として選定し中間画像を連続的に表示する機能を含む請求項 2 9 に記載の装置。

【請求項 3 5】 前記映像効果に関する機能は、処理の対象となる画像を二方向に補間して得られた中間画像を表示する機能を含む請求項 2 9 に記載の装置。

【請求項 3 6】 前記ユーティリティ機能は、生成すべき中間画像の枚数を指定する機能を含む請求項 3 0 に記載の装置。

【請求項 3 7】 前記生成すべき中間画像の枚数に関する指定は G U I を介してユーザから取得される請求項 3 6 に記載の装置。

【請求項 3 8】 前記モバイル機器への適用を前提とした機能は、第 1 画像または第 2 画像の一方のみと前記対応情報をもとに生成された中間画像を表示する機能である請求項 3 1 に記載の装置。

【請求項 3 9】 第 1 画像または第 2 画像の一方のみと前記対応情報をストーリーミング処理して前記中間画像が生成される請求項 3 8 に記載の装置。

【請求項 4 0】 前記中間画像、前記第 1 画像、前記第 2 画像のタイミング

を調整してそれらを表示する表示部をさらに含む請求項28から39のいずれかに記載の装置。

【請求項41】 第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第1画像と第2画像の中間画像を生成する工程と、

その中間画像の生成に当たり、電子商取引における商品プレゼンテーション機能を提供する工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【請求項42】 第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第1画像と第2画像の中間画像を生成する工程と、

その中間画像の生成に当たり、映像効果に関する機能を提供する工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【請求項43】 第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第1画像と第2画像の中間画像を生成する工程と、

その中間画像の生成に当たり、動画像の生成に関するユーティリティ機能を提供する工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【請求項44】 第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、

その対応情報もとに、第1画像と第2画像の中間画像を生成する工程と、

その中間画像の生成を、モバイル機器への適用を前提とした機能のもとで行う工程と、

をコンピュータに実行せしめることを特徴とするコンピュータプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

この発明は、映像処理技術に関し、とくにデジタル映像効果技術に関する。

【0002】

【従来の技術】

B S デジタル放送が開始のタイミングを迎え、また多数のユーザが P C や形態電話からインターネット上のサービスを享受するようになり、放送と通信の垣根が急速に崩れつつある。ブロードバンド時代においてマルチメディア文化は真の進展を遂げるであろうし、その中でもとりわけ動画の配信はキーテクノロジーとなる。人間が外界から取得する情報の量において、映像は音声の 2 0 0 倍の内容を伝達できるといわれる。映像は娯楽用途にとどまらず、広く人間の生活と文化を支えるソフトインフラとなる。映像がデジタル化されるにつれ、C G や画像処理技術の適用によって映像効果技術の活躍の場も広がる。

【0003】

【発明が解決しようとする課題】

従来より、いろいろな映像効果が提案され、利用されてきた。たとえば、ウォークスルー、多視点画像、特殊フィルタリング、擬似立体視などである。今後、映像の社会における重要性がさらに高まるにつれ、こうした映像効果がより高い品質をもって提供され、同時に伝送系の要請を満たすべく、高いデータ圧縮効果を伴うことが望まれる。この発明はそうした要望に対してなされたものであり、その目的のひとつは、動画の生成を比較的軽いデータで実現する映像効果技術の提供にある。本発明の別の目的は、映像効果に限らず、デジタルおよびインターネット時代における各種画像技術を実現する画像補間技術の提供にある。

【0004】

【課題を解決するための手段】

本発明のある態様は、第 1 画像上に設けられたメッシュの格子点に対応する第 2 画像上の点を画像マッチングによって検出する工程と、その結果をもとに、第 1 画像上のメッシュを構成する起点ポリゴンに、第 2 画像上にて対応する終点ポ

リゴンを定義する工程とを含む。

【0005】

検出する工程は、第1画像と第2画像の間で画素単位のマッチング計算を行ってもよい。この計算は、全画素について行ってもよいし、格子点のみ、または格子点とそれに関連する画素について行ってもよい。

【0006】

検出する工程は、第1画像について二次元的な探索を行って検出した特異点と、第2画像について二次元的な探索を行って検出した特異点との対応をもとに画素単位のマッチング計算を行ってもよい。この場合さらに、検出する工程は、第1画像と第2画像を、それぞれ特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得してもよい。

【0007】

定義する工程は、起点ポリゴンと終点ポリゴンの対応関係を対応点ファイルに出力し、このファイルを第1画像および第2画像と関連づけて保持してもよい。また本発明は、対応点ファイルを利用し、第1画像と第2画像の中間画像をポリゴン単位の補間計算にて生成する工程をさらに含んでもよい。

【0008】

前述の特異点を用いるマッチング方法は、本出願人が先に特許第2927350号にて提案した技術（以下「前提技術」という）の応用であり、前記検出する工程に好適である。ただし、前提技術は格子点またはそれによって決まるポリゴンに関する本発明の特徴にはいっさい触れていない。本発明でポリゴンという一種の簡略化技術を導入したため、前記対応点ファイルのサイズを非常に小さくすることができる。

【0009】

すなわち、第1、第2画像がそれぞれ $n \times m$ 画素を有する場合、それらの画素単位の対応をそのまま記述すると、 $(n \times m)^2$ という組合せが生じ、対応点ファイルが膨大なデータ量になる。しかし、これを格子点どうしの対応関係、また

は実質的に同じことであるが、格子点で決まるポリゴンどうしの対応関係を記述することに改めることでデータ量が大幅に減る。

【0010】

一方、第1、第2画像およびこの対応点ファイルがあれば、補間計算で中間画像が生成できる。したがって、本発明は映像効果でいうモーフィングを実現する。第1、第2画像をともし動画の中のフレームと考えれば、本発明はモーフィングではなく、動画の圧縮技術と把握できる。したがって、第1、第2画像および対応点ファイルのみを保有すれば、動画を再現でき、動画の伝送、ストレージ等において大きな効果を発揮する。

【0011】

本発明の別の態様は、映像効果装置に関する。この装置は、第1画像および第2画像を取得する画像入力部と、第1画像と第2画像のマッチングを計算するマッチングプロセッサとを含み、このマッチングプロセッサは、第1画像上のメッシュを構成する起点ポリゴンに、第2画像上にて対応する終点ポリゴンを定義することにより、対応点ファイルを生成する。

【0012】

マッチングプロセッサは、第1画像について二次元的な探索を行って検出した特異点と、第2画像について二次元的な探索を行って検出した特異点との対応をもとに画素単位のマッチング計算を行ってもよい。また、第1画像と第2画像を、それぞれ前記特異点を抽出することによって多重解像度化したうえで、同一解像度レベル間で画像単位のマッチング計算を行い、その結果を異なる解像度レベルにおけるマッチング計算に継承しながら、最終的に最も解像度の細かいレベルにおける画素単位の対応関係を取得してもよい。

【0013】

この装置はさらに、第1画像、第2画像およびそれらの間の対応点ファイルを取得し、それらふたつの画像の中間画像を生成する中間画像生成部を含み、前記対応点ファイルが第1画像上にとられたメッシュの格子点と、それに対応する第2画像上の点の位置関係を記述していてもよい。また、中間画像、第1画像、第2画像のタイミングを調整してそれらを表示する表示部をさらに含んでもよい。

なお、本発明は前提技術を必須とはしない。

【 0 0 1 4 】

本発明の別のさらに態様は、画像補間方法に関する。この方法は、第1画像上に設けられたメッシュを構成する起点ポリゴンと第2画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得する工程と、その対応情報もとに、第1画像と第2画像の中間画像を生成する工程と、その中間画像の生成に当たり、電子商取引における商品プレゼンテーション機能を提供する工程とを含む。

【 0 0 1 5 】

その機能に代えて、映像効果に関する機能、動画像の生成に関するユーティリティ機能を提供してもよい。また、中間画像の表示をモバイル機器への適用を前提とした機能のもとで行ってもよい。

【 0 0 1 6 】

商品プレゼンテーション機能は、ユーザからの指示にしたがって商品を回転して表示する機能を含んでもよい。また、商品画像を二方向、例えば左右と上下に補間して得られた中間画像を表示する機能を含んでもよい。

【 0 0 1 7 】

映像効果に関する機能は、3以上の画像の中から順次ふたつの画像を第1画像および第2画像として選定し中間画像を連続的に表示する機能を含んでもよい。例えば、多数の離散的な視点画像間をつぎつぎ補間して中間画像を生成すれば、あたかもパノラマ的な画像が得られる。視点を時間とともに動かしたかのごとく順に表示すれば、ウォークスルー画像が得られる。映像効果に関する機能は、処理の対象となる画像を二方向に補間して得られた中間画像を表示する機能を含んでもよい。

【 0 0 1 8 】

動画に関するユーティリティ機能は、生成すべき中間画像の枚数を指定する機能を含んでもよい。多数の中間画像を生成すれば、たとえば、スローモーション画像が生成できる。枚数を指定することで、動画の再生速度をインタラクティブに設定する途がひらかれる。生成すべき中間画像の枚数に関する指定はGUIを介してユーザから取得されてもよい。

【 0 0 1 9 】

モバイル機器への適用を前提とした機能は、第 1 画像または第 2 画像の一方のみと前記対応情報をもとに生成された中間画像を表示する機能であってもよい。その場合、画像をテンポラリにおくためのメモリの容量が少なくても済む。対応情報と第 1 画像のみで中間画像を生成する場合、第 1 画像の各画素その他の画像要素（以下単に画素で代表させる）を対応情報にしたがって移動させればよい。すなわち、通常は対応情報で画素の位置および画素値すなわち色の両方を補間できるが、これを画素の位置の移動のみの処理に簡素化する。もちろん、対応情報自体に第 2 画像の画素の色情報を組み込むことも可能であり、その場合は画素の位置と色の両方が良好に補間できる。

【 0 0 2 0 】

第 1 画像または第 2 画像の一方のみと前記対応情報をストリーミング処理して中間画像が生成されてもよい。この場合も、処理に必要なメモリの容量を減らす可能性があり、携帯電話その他の端末への実装に有利である。

【 0 0 2 1 】

本発明のさらに別の態様は、画像を補間する装置であって、第 1 画像上に設けられたメッシュを構成する起点ポリゴンと第 2 画像上に設けられたメッシュを構成する終点ポリゴンとの対応情報を取得し、その対応情報もとに第 1 画像と第 2 画像の中間画像を生成する中間画像生成部を含み、当該中間画像生成部は、中間画像の生成に当たり、電子商取引における商品プレゼンテーション機能を提供する。この他に、映像効果に関する機能、動画像の生成に関するユーティリティ機能、モバイル機器への適用を前提とした機能を提供してもよい。

【 0 0 2 2 】

この装置はさらに、中間画像、第 1 画像、第 2 画像のタイミングを調整してそれらを表示する表示部を含んでもよい。

【 0 0 2 3 】

なお、以上の各構成、工程を任意に入れ替えたり、方法と装置の間で表現を一部または全部入れ替え、または追加したり、表現をコンピュータプログラム、記録媒体等に変更したものもまた、本発明として有効である。

【 0 0 2 4 】

【発明の実施の形態】

はじめに、実施の形態で利用する多重解像度特異点フィルタ技術とそれを用いた画像マッチング処理を「前提技術」として詳述する。これらの技術は本出願人がすでに特許第 2 9 2 7 3 5 0 号を得ている技術であり、本発明との組合せに好適である。本発明では、画像上にメッシュを設け、その格子点によって多数の画素を代表されるため、もともと前提技術のような画素単位のマッチング技術に対する適用効果が高いためである。ただし、実施の形態で採用可能な画像マッチング技術はこれに限られない。

【 0 0 2 5 】

図 1 8 以降、前提技術を利用した画像補間技術および映像効果技術を具体的に説明する。

[前提技術の背景]

ふたつの画像の自動的なマッチング、つまり画像領域や画素どうしの対応付けは、コンピュータビジョンやコンピュータグラフィックスにおける最も難しくかつ重要なテーマのひとつである。例えば、あるオブジェクトに関して異なる視点からの画像間でマッチングがとれれば、他の視点からの画像を生成することができる。右目画像と左目画像のマッチングが計算できれば、立体画像を用いた写真測量も可能である。顔の画像のモデルと他の顔の画像のマッチングがとれたとき、目、鼻、口といった特徴的な顔の部分を抽出することができる。例えば人の顔と猫の顔の画像間でマッチングが正確にとられたとき、それらの中割画像を自動的に生成することでモーフィングを完全自動化することができる。

【 0 0 2 6 】

しかし従来一般に、ふたつの画像間の対応点は人がいちいち指定しなければならず、多大な作業工数を要した。この問題を解消するために数多くの対応点自動検出方法が提案されている。例えば、エピポーラ直線を用いることによって対応点の候補の数を減らす考えがある。しかし、その場合でも処理はきわめて複雑である。複雑さを低減するために、左目画像の各点の座標は通常右目画像でもほぼ同じ位置にあると想定される。しかし、こうした制約を設けると、大域的特徴及

び局所的特徴を同時に満たすマッチングをとることは非常に困難になる。

【0027】

ボリュームレンダリングでは、ボクセルを構成するために一連の断面画像が用いられる。この場合、従来一般に、上方の断面画像における画素が下方の断面画像の同一箇所にある画素と対応すると仮定され、これらの画素のペアが内挿計算に用いられる。このようにきわめて単純な方法を用いるため、連続する断面間の距離が遠く、オブジェクトの断面形状が大きく変化する場合、ボリュームレンダリングで構築されたオブジェクトは不明瞭になりがちである。

【0028】

立体写真測量法など、エッジの検出を利用するマッチングアルゴリズムも多い。しかしこの場合、結果的に得られる対応点の数が少ないため、マッチングのとれた対応点間のギャップを埋めるべく、ディスパリティの値を内挿計算しなければならない。一般にあらゆるエッジ検出器は、それらが用いる局所的なウィンドウの中で画素の輝度が変化したとき、これが本当にエッジの存在を示唆するかどうかを判断することが難しい。エッジ検出器は、本来的にすべてハイパスフィルタであり、エッジと同時にノイズも拾ってしまう。

【0029】

さらに別の手法として、オプティカルフローが知られている。二枚の画像が与えられたとき、オプティカルフローでは画像内のオブジェクト（剛体）の動きを検出する。その際、オブジェクトの各画素の輝度は変化しないと仮定する。オプティカルフローでは例えば (u, v) のベクトル場の滑らかさといった、いくつかの付加的な条件とともに、各画素の動きベクトル (u, v) を計算する。しかし、オプティカルフローでは画像間の大域的な対応関係を検出することはできない。画素の輝度の局所的な変化に注目するのみであり、画像の変位が大きい場合、システムの誤差は顕著になる。

【0030】

画像の大域的な構造を認識するために、多重解像度フィルタも数多く提案されてきた。それらは線形フィルタと非線形フィルタに分類される。前者の例としてウェーブレットがあるが、線形フィルタは一般に、画像マッチングにはさして有

用ではない。なぜなら、極値をとる画素の輝度に関する情報がそれらの位置情報とともに次第に不鮮明になるためである。図1(a)と図1(b)は顔の画像に対して平均化フィルタを適用した結果を示している。同図のごとく、極値をとる画素の輝度が平均化によって次第に薄れるとともに、位置も平均化の影響でシフトしていく。その結果、目(輝度の極小点)の輝度や位置の情報は、このような粗い解像度レベルで曖昧になり、この解像度では正しいマッチングを計算することができない。したがって、粗い解像度レベルを設けるのが大域的なマッチングのためでありながら、ここで得られたマッチングは画像の本当の特徴(目、つまり極小点)に正確に対応しない。より精細な解像度レベルで目が鮮明に現れたとしても、大域的なマッチングをとる際に混入した誤差は、もはや取り返しがつかない。入力画像にスムージング処理を加えることにより、テクスチャ領域のステレオ情報が落ちてしまうこともすでに指摘されている。

【0031】

一方、最近地形学の分野で利用されはじめた非線形フィルタとして一次元の「ふるい(sieve)」演算子がある。この演算子は、所定の大きさの一次元ウィンドウ内の極小値(または極大値)を選択することにより、縮尺と空間の因果関係を保存しながら画像にスムージング処理を加える。その結果得られる画像は元の画像と同じ大きさであるが、小さな波の成分が取り除かれているため、より単純になる。画像の情報を落とすという点で、この演算子は広い意味での「多重解像度フィルタ」に分類することはできるが、実際にはウェーブレットのように画像の解像度を変えながら画像を階層化するわけではなく(つまり狭い意味での多重解像度フィルタではなく)、画像間の対応の検出には利用できない。

【0032】

[前提技術が解決しようとする課題]

以上をまとめれば以下の課題が認められる。

1. 画像の特徴を正確に、かつ比較的簡単な処理で把握する画像処理方法が乏しかった。特に、特徴のある点に関する情報、例えば画素値や位置を維持しながら特徴を抽出できる画像処理方法に関する有効な提案が少なかった。
2. 画像の特徴をもとに対応点を自動検出する場合、一般に処理が複雑である

か、ノイズ耐性が低いなどの欠点があった。また、処理に際していろいろな制約を設ける必要があり、大域的特徴及び局所的特徴を同時に満たすマッチングをとることが困難だった。

3. 画像の大域的な構造または特徴を認識するために多重解像度フィルタを導入しても、そのフィルタが線形フィルタの場合、画素の輝度情報と位置情報が曖昧になった。その結果、対応点の把握が不正確になりやすかった。非線形フィルタである一次元ふるい演算子は画像を階層化しないため、画像間の対応点の検出には利用できなかった。

4. これらの結果、対応点を正しく把握しようとするれば、結局人手による指定に頼るほか有効な手だてがなかった。

【0033】

前提技術はこれらの課題の解決を目的としてなされたものであり、画像処理の分野において、画像の特徴の的確な把握を可能にする技術を提供するものである。

【0034】

[前提技術が課題を解決するための手段]

この目的のために前提技術のある態様は、新たな多重解像度の画像フィルタを提案する。この多重解像度フィルタは画像から特異点を抽出する。したがって、特異点フィルタともよばれる。特異点とは画像上特徴をもつ点をいう。例として、ある領域において画素値（画素値とは、色番号、輝度値など画像または画素に関する任意の数値を指す）が最大になる極大点、最小になる極小点、ある方向については最大だが別の方向については最小になるような鞍点がある。特異点は位相幾何学上の概念であってもよい。ただし、その他どのような特徴を有してもよい。いかなる性質の点を特異点と考えるかは、前提技術にとって本質問題ではない。

【0035】

この態様では、多重解像度フィルタを用いた画像処理が行われる。まず検出工程において、第一の画像に対し、二次元的な探索を行って特異点が検出される。つぎに生成工程において、検出された特異点を抽出して第一の画像よりも解像度

の低い第二の画像が生成される。第二の画像には第一の画像のもつ特異点が引き継がれる。第二の画像は第一の画像よりも解像度が低いため、画像の大域的な特徴の把握に好適である。

【0036】

前提技術の別の態様は特異点フィルタを用いた画像マッチング方法に関する。この態様では、始点画像と終点画像間のマッチングがとられる。始点画像および終点画像とは、ふたつの画像の区別のために便宜的に与えた名称であり、本質的な違いはない。

【0037】

この態様では、第一工程にて、始点画像に特異点フィルタを施して解像度の異なる一連の始点階層画像が生成される。第二工程では、終点画像に特異点フィルタを施して解像度の異なる一連の終点階層画像が生成される。始点階層画像、終点階層画像とは、それぞれ始点画像、終点画像を階層化して得られる画像群をいい、それぞれ最低2枚の画像からなる。つぎに第三工程において、始点階層画像と終点階層画像のマッチングが解像度レベルの階層の中で計算される。この態様によれば、多重解像度フィルタによって特異点に関連する画像の特徴が抽出され、および／または明確化されるため、マッチングが容易になる。マッチングのための拘束条件は特に必要としない。

【0038】

前提技術のさらに別の態様も始点画像と終点画像のマッチングに関する。この態様では、予め複数のマッチング評価項目のそれぞれに関して評価式を設け、それらの評価式を統合して総合評価式を定義し、その総合評価式の極値付近に注目して最適マッチングを探索する。総合評価式は、評価式の少なくともひとつに係数パラメータを掛けたうえでそれらの評価式の総和として定義してもよく、その場合、総合評価式またはいずれかの評価式がほぼ極値をとる状態を検出して前記パラメータを決定してもよい。「極値付近」または「ほぼ極値をとる」としたのは、多少誤差を含んでいてもよいためである。多少の誤差は前提技術にはさして問題とならない。

【0039】

極値自体も前記パラメータに依存するため、極値の挙動、つまり極値の変化の様子をもとに、最適と考えられるパラメータを決定する余地が生じる。この態様はその事実を利用している。この態様によれば、元来調整の困難なパラメータの決定を自動化する途が拓かれる。

【0040】

[前提技術の実施の形態]

最初に[1]で前提技術の要素技術の詳述し、[2]で処理手順を具体的に説明する。さらに[3]で実験の結果を報告する。

【0041】

[1] 要素技術の詳細

[1. 1] イントロダクション

特異点フィルタと呼ばれる新たな多重解像度フィルタを導入し、画像間のマッチングを正確に計算する。オブジェクトに関する予備知識は一切不要である。画像間のマッチングの計算は、解像度の階層を進む間、各解像度において計算される。その際、粗いレベルから精細なレベルへと順に解像度の階層を辿っていく。計算に必要なパラメータは、人間の視覚システムに似た動的計算によって完全に自動設定される。画像間の対応点を人手で特定する必要はない。

【0042】

本前提技術は、例えば完全に自動的なモーフィング、物体認識、立体写真測量、ボリュームレンダリング、少ないフレームからの滑らかな動画の生成などに応用できる。モーフィングに用いる場合、与えられた画像を自動的に変形することができる。ボリュームレンダリングに用いる場合、断面間の中間的な画像を正確に再構築することができる。断面間の距離が遠く、断面の形状が大きく変化する場合でも同様である。

【0043】

[1. 2] 特異点フィルタの階層

前提技術に係る多重解像度特異点フィルタは、画像の解像度を落としながら、しかも画像に含まれる各特異点の輝度及び位置を保存することができる。ここで画像の幅を N 、高さを M とする。以下簡単のため、 $N=M=2^n$ (n は自然数)

と仮定する。また、区間 $[0, N] \subset R$ を I と記述する。 (i, j) における画像の画素を $p(i, j)$ と記述する ($i, j \in I$)。

【0044】

ここで多重解像度の階層を導入する。階層化された画像群は多重解像度フィルタで生成される。多重解像度フィルタは、もとの画像に対して二次元的な探索を行って特異点を検出し、検出された特異点を抽出してもとの画像よりも解像度の低い別の画像を生成する。ここで第 m レベルにおける各画像のサイズは $2^m \times 2^m$ ($0 \leq m \leq n$) とする。特異点フィルタは次の4種類の新たな階層画像を n から下がる方向で再帰的に構築する。

【0045】

【数1】

$$\begin{aligned} p_{(i,j)}^{(m,0)} &= \min(\min(p_{(2i,2j)}^{(m+1,0)}, p_{(2i,2j+1)}^{(m+1,0)}), \min(p_{(2i+1,2j)}^{(m+1,0)}, p_{(2i+1,2j+1)}^{(m+1,0)})) \\ p_{(i,j)}^{(m,1)} &= \max(\min(p_{(2i,2j)}^{(m+1,1)}, p_{(2i,2j+1)}^{(m+1,1)}), \min(p_{(2i+1,2j)}^{(m+1,1)}, p_{(2i+1,2j+1)}^{(m+1,1)})) \\ p_{(i,j)}^{(m,2)} &= \min(\max(p_{(2i,2j)}^{(m+1,2)}, p_{(2i,2j+1)}^{(m+1,2)}), \max(p_{(2i+1,2j)}^{(m+1,2)}, p_{(2i+1,2j+1)}^{(m+1,2)})) \\ p_{(i,j)}^{(m,3)} &= \max(\max(p_{(2i,2j)}^{(m+1,3)}, p_{(2i,2j+1)}^{(m+1,3)}), \max(p_{(2i+1,2j)}^{(m+1,3)}, p_{(2i+1,2j+1)}^{(m+1,3)})) \end{aligned}$$

(式1)

ただしここで、

【数2】

$$p_{(i,j)}^{(n,0)} = p_{(i,j)}^{(n,1)} = p_{(i,j)}^{(n,2)} = p_{(i,j)}^{(n,3)} = p_{(i,j)} \quad (\text{式2})$$

とする。以降これら4つの画像を副画像（サブイメージ）と呼ぶ。 $\min_{x \leq t \leq x+1}$ 、 $\max_{x \leq t \leq x+1}$ をそれぞれ α 及び β と記述すると、副画像はそれぞれ以下のように記述できる。

【0046】

$$\begin{aligned} p(m, 0) &= \alpha(x) \alpha(y) p(m+1, 0) \\ p(m, 1) &= \alpha(x) \beta(y) p(m+1, 1) \\ p(m, 2) &= \beta(x) \alpha(y) p(m+1, 2) \end{aligned}$$

$$p(m, 3) = \beta(x) \beta(y) p(m+1, 3)$$

すなわち、これらは α と β のテンソル積のようなものと考えられる。副画像はそれぞれ特異点に対応している。これらの式から明らかなように、特異点フィルタはもとの画像について 2×2 画素で構成されるブロックごとに特異点を検出する。その際、各ブロックのふたつの方向、つまり縦と横について、最大画素値または最小画素値をもつ点を探索する。画素値として、前提技術では輝度を採用するが、画像に関するいろいろな数値を採用することができる。ふたつの方向の両方について最大画素値となる画素は極大点、ふたつの方向の両方について最小画素値となる画素は極小点、ふたつの方向の一方について最大画素値となるとともに、他方について最小画素値となる画素は鞍点として検出される。

【0047】

特異点フィルタは、各ブロックの内部で検出された特異点の画像（ここでは1画素）でそのブロックの画像（ここでは4画素）を代表させることにより、画像の解像度を落とす。特異点の理論的な観点からすれば、 $\alpha(x) \alpha(y)$ は極小点を保存し、 $\beta(x) \beta(y)$ は極大点を保存し、 $\alpha(x) \beta(y)$ 及び $\beta(x) \alpha(y)$ は鞍点を保存する。

【0048】

はじめに、マッチングをとるべき始点（ソース）画像と終点（デスティネーション）画像に対して別々に特異点フィルタ処理を施し、それぞれ一連の画像群、すなわち始点階層画像と終点階層画像を生成しておく。始点階層画像と終点階層画像は、特異点の種類に対応してそれぞれ4種類ずつ生成される。

【0049】

この後、一連の解像度レベルの中で始点階層画像と終点階層画像のマッチングがとれていく。まず $p(m, 0)$ を用いて極小点のマッチングがとられる。次に、その結果に基づき、 $p(m, 1)$ を用いて鞍点のマッチングがとられ、 $p(m, 2)$ を用いて他の鞍点のマッチングがとられる。そして最後に $p(m, 3)$ を用いて極大点のマッチングがとられる。

【0050】

図1(c)と図1(d)はそれぞれ図1(a)と図1(b)の副画像 $p(5,$

0)を示している。同様に、図1(e)と図1(f)は $p(5, 1)$ 、図1(g)と図1(h)は $p(5, 2)$ 、図1(i)と図1(j)は $p(5, 3)$ をそれぞれ示している。これらの図からわかるとおり、副画像によれば画像の特徴部分のマッチングが容易になる。まず $p(5, 0)$ によって目が明確になる。目は顔の中で輝度の極小点だからである。 $p(5, 1)$ によれば口が明確になる。口は横方向で輝度が低いためである。 $p(5, 2)$ によれば首の両側の縦線が明確になる。最後に、 $p(5, 3)$ によって耳や頬の最も明るい点が明確になる。これらは輝度の極大点だからである。

【0051】

特異点フィルタによれば画像の特徴が抽出できるため、例えばカメラで撮影された画像の特徴と、予め記録しておいたいくつかのオブジェクトの特徴を比較することにより、カメラに映った被写体を識別することができる。

【0052】

[1. 3] 画像間の写像の計算

始点画像の位置 (i, j) の画素を $p^{(n)}(i, j)$ と書き、同じく終点画像の位置 (k, l) の画素を $q^{(n)}(k, l)$ で記述する。 $i, j, k, l \in I$ とする。画像間の写像のエネルギー(後述)を定義する。このエネルギーは、始点画像の画素の輝度と終点画像の対応する画素の輝度の差、及び写像の滑らかさによって決まる。最初に最小のエネルギーを持つ $p(m, 0)$ と $q(m, 0)$ 間の写像 $f(m, 0) : p(m, 0) \rightarrow q(m, 0)$ が計算される。 $f(m, 0)$ に基づき、最小エネルギーを持つ $p(m, 1)$ 、 $q(m, 1)$ 間の写像 $f(m, 1)$ が計算される。この手続は、 $p(m, 3)$ と $q(m, 3)$ の間の写像 $f(m, 3)$ の計算が終了するまで続く。各写像 $f(m, i)$ ($i = 0, 1, 2, \dots$)を副写像と呼ぶことにする。 $f(m, i)$ の計算の都合のために、 i の順序は次式のように並べ替えることができる。並べ替えが必要な理由は後述する。

【0053】

【数3】

$$f^{(m,i)} : p^{(m,\sigma(i))} \rightarrow q^{(m,\sigma(i))} \quad (\text{式3})$$

ここで $\sigma(i) \in \{0, 1, 2, 3\}$ である。

【0054】

[1. 3. 1] 全単射

始点画像と終点画像の間のマッチングを写像で表現する場合、その写像は両画像間で全単射条件を満たすべきである。両画像に概念上の優劣はなく、互いの画素が全射かつ単射で接続されるべきだからである。しかしながら通常の場合とは異なり、ここで構築すべき写像は全単射のデジタル版である。前提技術では、画素は格子点によって特定される。

【0055】

始点副画像（始点画像について設けられた副画像）から終点副画像（終点画像について設けられた副画像）への写像は、 $f^{(m, s)} : I/2^{n-m} \times I/2^{n-m} \rightarrow I/2^{n-m} \times I/2^{n-m}$ ($s = 0, 1, \dots$) によって表される。ここで、 $f^{(m, s)}(i, j) = (k, l)$ は、始点画像の $p^{(m, s)}(i, j)$ が終点画像の $q^{(m, s)}(k, l)$ に写像されることを意味する。簡単のために、 $f(i, j) = (k, l)$ が成り立つとき画素 $q(k, l)$ を $q_f(i, j)$ と記述する。

【0056】

前提技術で扱う画素（格子点）のようにデータが離散的な場合、全単射の定義は重要である。ここでは以下のように定義する（ i, i', j, j', k, l は全て整数とする）。まず始めに、始点画像の平面において R によって表記される各正方形領域、

【数4】

$$p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}$$

(式4)

を考える（ $i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$ ）。ここで R の各辺（エッジ）の方向を以下のように定める。

【0057】

【数 5】

$$\overrightarrow{p_{(i,j)}^{(m,s)} p_{(i+1,j)}^{(m,s)}} \rightarrow \overrightarrow{p_{(i+1,j)}^{(m,s)} p_{(i+1,j+1)}^{(m,s)}} \rightarrow \overrightarrow{p_{(i+1,j+1)}^{(m,s)} p_{(i,j+1)}^{(m,s)}} \text{ and } \overrightarrow{p_{(i,j+1)}^{(m,s)} p_{(i,j)}^{(m,s)}} \quad (\text{式 5})$$

この正方形は写像 f によって終点画像平面における四辺形に写像されなければならない。 $f(m, s)(R)$ によって示される四辺形、

【数 6】

$$q_f^{(m,s)}(i,j) q_f^{(m,s)}(i+1,j) q_f^{(m,s)}(i+1,j+1) q_f^{(m,s)}(i,j+1) \quad (\text{式 6})$$

は、以下の全単射条件を満たす必要がある。

【0058】

1. 四辺形 $f(m, s)(R)$ のエッジは互いに交差しない。
2. $f(m, s)(R)$ のエッジの方向は R のそれらに等しい (図 2 の場合、時計回り)。
3. 緩和条件として収縮写像 (リトラクション: retractions) を許す。

【0059】

何らかの緩和条件を設けないかぎり、全単射条件を完全に満たす写像は単位写像しかないためである。ここでは $f(m, s)(R)$ のひとつのエッジの長さが 0、すなわち $f(m, s)(R)$ は三角形になってもよい。しかし、面積が 0 となるような図形、すなわち 1 点または 1 本の線分になってはならない。図 2 (R) がもとの四辺形の場合、図 2 (A) と図 2 (D) は全単射条件を満たすが、図 2 (B)、図 2 (C)、図 2 (E) は満たさない。

【0060】

実際のインプリメンテーションでは、写像が全射であることを容易に保証すべく、さらに以下の条件を課してもよい。つまり始点画像の境界上の各画素は、終点画像において同じ位置を占める画素に写影されるというものである。すなわち、 $f(i, j) = (i, j)$ (ただし $i = 0, i = 2^m - 1, j = 0, j = 2^m - 1$ の 4 本の線上) である。この条件を以下「付加条件」とも呼ぶ。

【0061】

[1. 3. 2] 写像のエネルギー

[1. 3. 2. 1] 画素の輝度に関するコスト

写像 f のエネルギーを定義する。エネルギーが最小になる写像を探すことが目的である。エネルギーは主に、始点画像の画素の輝度とそれに対応する終点画像の画素の輝度の差で決まる。すなわち、写像 $f^{(m, s)}$ の点 (i, j) におけるエネルギー $C^{(m, s)}(i, j)$ は次式によって定まる。

【0062】

【数7】

$$C_{(i,j)}^{(m,s)} = |V(p_{(i,j)}^{(m,s)}) - V(q_{f(i,j)}^{(m,s)})|^2 \quad (\text{式7})$$

ここで、 $V(p^{(m, s)}(i, j))$ 及び $V(q^{(m, s)}_{f(i, j)})$ はそれぞれ画素 $p^{(m, s)}(i, j)$ 及び $q^{(m, s)}_{f(i, j)}$ の輝度である。 f のトータルのエネルギー $C^{(m, s)}$ は、マッチングを評価するひとつの評価式であり、つぎに示す $C^{(m, s)}(i, j)$ の合計で定義できる。

【0063】

【数8】

$$C_f^{(m,s)} = \sum_{i=0}^{2^m-1} \sum_{j=0}^{2^m-1} C_{(i,j)}^{(m,s)} \quad (\text{式8})$$

[1. 3. 2. 2] 滑らかな写像のための画素の位置に関するコスト

滑らかな写像を得るために、写像に関する別のエネルギー D_f を導入する。このエネルギーは画素の輝度とは関係なく、 $p^{(m, s)}(i, j)$ および $q^{(m, s)}_{f(i, j)}$ の位置によって決まる ($i = 0, \dots, 2^m - 1, j = 0, \dots, 2^m - 1$)。点 (i, j) における写像 $f^{(m, s)}$ のエネルギー $D^{(m, s)}(i, j)$ は次式で定義される。

【0064】

【数9】

$$D_{(i,j)}^{(m,s)} = \eta E_{0(i,j)}^{(m,s)} + E_{1(i,j)}^{(m,s)} \quad (\text{式9})$$

ただし、係数パラメータ η は0以上の実数であり、また、

【数 10】

$$E_{0(i,j)}^{(m,s)} = \|(i,j) - f^{(m,s)}(i,j)\|^2 \quad (\text{式 10})$$

【数 11】

$$E_{1(i,j)}^{(m,s)} = \sum_{i'=i-1}^i \sum_{j'=j-1}^j \|(f^{(m,s)}(i,j) - (i,j)) - (f^{(m,s)}(i',j') - (i',j'))\|^2 / 4 \quad (\text{式 11})$$

とする。ここで、

【数 12】

$$\|(x,y)\| = \sqrt{x^2 + y^2} \quad (\text{式 12})$$

であり、 $i' < 0$ および $j' < 0$ に対して $f(i', j')$ は 0 と決める。 E_0 は (i, j) 及び $f(i, j)$ の距離で決まる。 E_0 は画素があまりにも離れた画素へ写影されることを防ぐ。ただし E_0 は、後に別のエネルギー関数で置き換える。 E_1 は写像の滑らかさを保証する。 E_1 は、 $p(i, j)$ の変位とその隣接点の変位の間の隔たりを表す。以上の考察をもとに、マッチングを評価する別の評価式であるエネルギー D_f は次式で定まる。

【0065】

【数 13】

$$D_f^{(m,s)} = \sum_{i=0}^{i=2^m-1} \sum_{j=0}^{j=2^m-1} D_{(i,j)}^{(m,s)} \quad (\text{式 13})$$

[1. 3. 2. 3] 写像の総エネルギー

写像の総エネルギー、すなわち複数の評価式の統合に係る総合評価式は $\lambda C(m, s)_f + D(m, s)_f$ で定義される。ここで係数パラメータ λ は 0 以上の実数である。目的は総合評価式が極値をとる状態を検出すること、すなわち次式で示す最小エネルギーを与える写像を見いだすことである。

【0066】

【数 14】

$$\min_f \lambda C_f^{(m,s)} + D_f^{(m,s)} \quad (\text{式 14})$$

$\lambda = 0$ 及び $\eta = 0$ の場合、写像は単位写像になることに注意すべきである（す

なわち、全ての $i = 0, \dots, 2^m - 1$ 及び $j = 0, \dots, 2^m - 1$ に対して $f^{(m, s)}(i, j) = (i, j)$ となる)。後述のごとく、本前提技術では最初に $\lambda = 0$ 及び $\eta = 0$ の場合を評価するため、写像を単位写像から徐々に変形していくことができる。仮に総合評価式の λ の位置を変えて $C^{(m, s)}_f + \lambda D^{(m, s)}_f$ と定義したとすれば、 $\lambda = 0$ 及び $\eta = 0$ の場合に総合評価式が $C^{(m, s)}_f$ だけになり、本来何等関連のない画素どうしが単に輝度が近いというだけで対応づけられ、写像が無意味なものになる。そうした無意味な写像をもとに写像を変形していてもまったく意味をなさない。このため、単位写像が評価の開始時点で最良の写像として選択されるよう係数パラメータの与えかたが配慮されている。

【0067】

オブティカルフローもこの前提技術同様、画素の輝度の差と滑らかさを考慮する。しかし、オブティカルフローは画像の変換に用いることはできない。オブジェクトの局所的な動きしか考慮しないためである。前提技術に係る特異点フィルタを用いることによって大域的な対応関係を検出することができる。

【0068】

[1. 3. 3] 多重解像度の導入による写像の決定

最小エネルギーを与え、全単射条件を満足する写像 f_{min} を多重解像度の階層を用いて求める。各解像度レベルにおいて始点副画像及び終点副画像間の写像を計算する。解像度の階層の最上位（最も粗いレベル）からスタートし、各解像度レベルの写像を、他のレベルの写像を考慮に入れながら決定する。各レベルにおける写像の候補の数は、より高い、つまりより粗いレベルの写像を用いることによって制限される。より具体的には、あるレベルにおける写像の決定に際し、それよりひとつ粗いレベルにおいて求められた写像が一種の拘束条件として課される。

【0069】

まず、

【数15】

$$(i', j') = ([\frac{i}{2}], [\frac{j}{2}]) \quad (\text{式15})$$

が成り立つとき、 $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ をそれぞれ $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ の parent と呼ぶことにする。[x] は x を越えない最大整数である。また $p^{(m, s)}(i, j)$ 、 $q^{(m, s)}(i, j)$ をそれぞれ $p^{(m-1, s)}(i', j')$ 、 $q^{(m-1, s)}(i', j')$ の child と呼ぶ。関数 parent (i, j) は次式で定義される。

【0070】

【数16】

$$parent(i, j) = ([\frac{i}{2}], [\frac{j}{2}]) \quad (式16)$$

$p^{(m, s)}(i, j)$ と $q^{(m, s)}(k, 1)$ の間の写像 $f^{(m, s)}$ は、エネルギー計算を行って最小になったものを見つけることで決定される。 $f^{(m, s)}(i, j) = (k, 1)$ の値は $f^{(m-1, s)}(m=1, 2, \dots, n)$ を用いることによって、以下のように決定される。まず、 $q^{(m, s)}(k, 1)$ は次の四辺形の内部になければならないという条件を課し、全単射条件を満たす写像のうち現実性の高いものを絞り込む。

【0071】

【数17】

$$q_{g^{(m, s)}(i-1, j-1)}^{(m, s)} q_{g^{(m, s)}(i-1, j+1)}^{(m, s)} q_{g^{(m, s)}(i+1, j+1)}^{(m, s)} q_{g^{(m, s)}(i+1, j-1)}^{(m, s)} \quad (式17)$$

ただしここで、

【数18】

$$g^{(m, s)}(i, j) = f^{(m-1, s)}(parent(i, j)) + f^{(m-1, s)}(parent(i, j) + (1, 1)) \quad (式18)$$

である。こうして定めた四辺形を、以下 $p^{(m, s)}(i, j)$ の相続 (inherited) 四辺形と呼ぶことにする。相続四辺形の内部において、エネルギーを最小にする画素を求める。

【0072】

図3は以上の手順を示している。同図において、始点画像のA, B, C, Dの画素は、第 $m-1$ レベルにおいてそれぞれ終点画像のA', B', C', D'へ写影される。画素 $p^{(m, s)}(i, j)$ は、相続四辺形A' B' C' D'の内部に存在する画素 $q^{(m, s)}(i, j)$ へ写影されなければならない。以上の配慮により、第 $m-1$ レベルの写像から第 m レベルの写像への橋渡しがなされる。

【0073】

先に定義したエネルギー E_0 は、第 m レベルにおける副写像 $f^{(m, 0)}$ を計算するために、次式に置き換える。

【0074】

【数19】

$$E_{0(i,j)} = \|f^{(m,0)}(i,j) - g^{(m)}(i,j)\|^2 \quad (\text{式19})$$

また、副写像 $f^{(m, s)}$ を計算するためには次式を用いる。

【0075】

【数20】

$$E_{0(i,j)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2 \quad (1 \leq i) \quad (\text{式20})$$

こうしてすべての副写像のエネルギーを低い値に保つ写像が得られる。式20により、異なる特異点に対応する副写像が、副写像どうしの類似度が高くなるように同一レベル内で関連づけられる。式19は、 $f^{(m, s)}(i, j)$ と、第 $m-1$ レベルの画素の一部と考えた場合の (i, j) が射影されるべき点の位置との距離を示している。

【0076】

仮に、相続四辺形A' B' C' D'の内部に全単射条件を満たす画素が存在しない場合は以下の措置をとる。まず、A' B' C' D'の境界線からの距離がL(始めは $L=1$)である画素を調べる。それらのうち、エネルギーが最小になるものが全単射条件を満たせば、これを $f^{(m, s)}(i, j)$ の値として選択する。そのような点が発見されるか、またはLがその上限の $L^{(m)}_{\max}$ に到達するまで、Lを大きくしていく。 $L^{(m)}_{\max}$ は各レベル m に対して固定であ

る。そのような点が全く発見されない場合、全単射の第3の条件を一時的に無視して変換先の四辺形の面積がゼロになるような写像も認め、 $f(m, s)(i, j)$ を決定する。それでも条件を満たす点が見つからない場合、つぎに全単射の第1及び第2条件を外す。

【0077】

多重解像度を用いる近似法は、写像が画像の細部に影響されることを回避しつつ、画像間の大域的な対応関係を決定するために必須である。多重解像度による近似法を用いなければ、距離の遠い画素間の対応関係を見いだすことは不可能である。その場合、画像のサイズはきわめて小さなものに限定しなければならず、変化の小さな画像しか扱うことができない。さらに、通常写像に滑らかさを要求するため、そうした画素間の対応関係を見つけにくくしている。距離のある画素から画素への写像のエネルギーは高いためである。多重解像度を用いた近似法によれば、そうした画素間の適切な対応関係を見いだすことができる。それらの距離は、解像度の階層の上位レベル（粗いレベル）において小さいためである。

【0078】

[1. 4] 最適なパラメータ値の自動決定

既存のマッチング技術の主な欠点のひとつに、パラメータ調整の困難さがある。大抵の場合、パラメータの調整は人手作業によって行われ、最適な値を選択することはきわめて難しい。前提技術に係る方法によれば、最適なパラメータ値を完全に自動決定することができる。

【0079】

前提技術に係るシステムはふたつのパラメータ、 λ 及び η を含む。端的にいえば、 λ は画素の輝度の差の重みであり、 η は写像の剛性を示している。これらのパラメータの値は初期値が0であり、まず $\eta = 0$ に固定して λ を0から徐々に増加させる。 λ の値を大きくしながら、しかも総合評価式（式14）の値を最小にする場合、各副写像に関する $C(m, s)_f$ の値は一般に小さくなっていく。このことは基本的にふたつの画像がよりマッチしなければならないことを意味する。しかし、 λ が最適値を超えると以下の現象が発生する。

【0080】

1. 本来対応すべきではない画素どうしが、単に輝度が近いというだけで誤って対応づけられる。

2. その結果、画素どうしの対応関係がおかしくなり、写像がくずれはじめる。

【0081】

3. その結果、式14において $D(m, s)_f$ が急激に増加しようとする。

4. その結果、式14の値が急激に増加しようとするため、 $D(m, s)_f$ の急激な増加を抑制するよう $f(m, s)$ が変化し、その結果 $C(m, s)_f$ が増加する。

したがって、 λ を増加させながら式14が最小値をとるという状態を維持しつつ $C(m, s)_f$ が減少から増加に転じる閾値を検出し、その λ を $\eta = 0$ における最適値とする。つぎに η を少しずつ増やして $C(m, s)_f$ の挙動を検査し、後述の方法で η を自動決定する。その η に対応して λ も決まる。

【0082】

この方法は、人間の視覚システムの焦点機構の動作に似ている。人間の視覚システムでは、一方の目を動かしながら左右両目の画像のマッチングがとられる。オブジェクトがはっきりと認識できるとき、その目が固定される。

【0083】

[1. 4. 1] λ の動的決定

λ は0から所定の刻み幅で増加されていき、 λ の値が変わる度に副写像が評価される。式14のごとく、総エネルギーは $\lambda C(m, s)_f + D(m, s)_f$ によって定義される。式9の $D(m, s)_f$ は滑らかさを表すもので、理論的には単位写像の場合に最小になり、写像が歪むほど E_0 も E_1 も増加していく。 E_1 は整数であるから、 $D(m, s)_f$ の最小刻み幅は1である。このため、現在の $\lambda C(m, s)_{(i, j)}$ の変化(減少量)が1以上でなければ、写像を変化させることによって総エネルギーを減らすことはできない。なぜなら、写像の変化に伴って $D(m, s)_f$ は1以上増加するため、 $\lambda C(m, s)_{(i, j)}$ が1以上減少しない限り総エネルギーは減らないためである。

【0084】

この条件のもと、 λ の増加に伴い、正常な場合に $C^{(m, s)}(i, j)$ が減少することを示す。 $C^{(m, s)}(i, j)$ のヒストグラムを $h(l)$ と記述する。 $h(l)$ はエネルギー $C^{(m, s)}(i, j)$ が l^2 である画素の数である。 $\lambda_1^2 \geq 1$ が成り立つために、例えば $l^2 = 1/\lambda$ の場合を考える。 λ が λ_1 から λ_2 まで微小量変化するとき、

【数21】

$$A = \sum_{l=\lceil \frac{1}{\lambda_2} \rceil}^{\lceil \frac{1}{\lambda_1} \rceil} h(l) \simeq \int_{\frac{1}{\lambda_2}}^{\frac{1}{\lambda_1}} h(l) dl = - \int_{\lambda_2}^{\lambda_1} h(l) \frac{1}{\lambda^{3/2}} d\lambda = \int_{\lambda_1}^{\lambda_2} \frac{h(l)}{\lambda^{3/2}} d\lambda$$

(式21)

で示されるA個の画素が、

【数22】

$$C_f^{(m, s)} - l^2 = C_f^{(m, s)} - \frac{1}{\lambda} \quad (\text{式22})$$

のエネルギーを持つより安定的な状態に変化する。ここでは仮に、これらの画素のエネルギーがすべてゼロになると近似している。この式は $C^{(m, s)}_f$ の値が、

【数23】

$$\partial C_f^{(m, s)} = -\frac{A}{\lambda} \quad (\text{式23})$$

だけ変化することを示し、その結果、

【数24】

$$\frac{\partial C_f^{(m, s)}}{\partial \lambda} = -\frac{h(l)}{\lambda^{5/2}} \quad (\text{式24})$$

が成立する。 $h(l) > 0$ であるから、通常 $C^{(m, s)}_f$ は減少する。しかし、 λ が最適値を越えようとするとき、上述の現象、つまり $C^{(m, s)}_f$ の増加が発生する。この現象を検出することにより、 λ の最適値を決定する。

【0085】

なお、 $H(h > 0)$ 及び k を定数とすると、

【数 2 5】

$$h(l) = H l^k = \frac{H}{\lambda^{k/2}} \quad (\text{式 2 5})$$

と仮定すれば、

【数 2 6】

$$\frac{\partial C_f^{(m,s)}}{\partial \lambda} = -\frac{H}{\lambda^{5/2+k/2}} \quad (\text{式 2 6})$$

が成り立つ。このとき $k \neq -3$ であれば、

【数 2 7】

$$C_f^{(m,s)} = C + \frac{H}{(3/2 + k/2)\lambda^{3/2+k/2}} \quad (\text{式 2 7})$$

となる。これが $C^{(m, s)}_f$ の一般式である (C は定数)。

【0086】

λ の最適値を検出する際、さらに安全を見て、全単射条件を破る画素の数を検査してもよい。ここで各画素の写像を決定する際、全単射条件を破る確率を p_0 と仮定する。この場合、

【数 2 8】

$$\frac{\partial A}{\partial \lambda} = \frac{h(l)}{\lambda^{3/2}} \quad (\text{式 2 8})$$

が成立しているため、全単射条件を破る画素の数は次式の率で増加する。

【0087】

【数 2 9】

$$B_0 = \frac{h(l)p_0}{\lambda^{3/2}} \quad (\text{式 2 9})$$

従って、

【数 3 0】

$$\frac{B_0 \lambda^{3/2}}{p_0 h(l)} = 1 \quad (\text{式 3 0})$$

は定数である。仮に $h(1) = H 1^k$ を仮定するとき、例えば、

【数 31】

$$B_0 \lambda^{3/2+k/2} = p_0 H \quad (\text{式 31})$$

は定数になる。しかし λ が最適値を越えると、上の値は急速に増加する。この現象を検出し、 $B_0 \lambda^{3/2+k/2} / 2^m$ の値が異常値 $B_{0 \text{ thresh}}$ を越えるかどうかを検査し、 λ の最適値を決定することができる。同様に、 $B_1 \lambda^{3/2+k/2} / 2^m$ の値が異常値 $B_{1 \text{ thresh}}$ を越えるかどうかを検査することにより、全単射の第3の条件を破る画素の増加率 B_1 を確認する。ファクター 2^m を導入する理由は後述する。このシステムはこれら2つの閾値に敏感ではない。これらの閾値は、エネルギー $C(m, s)_f$ の観察では検出し損なった写像の過度の歪みを検出するために用いることができる。

【0088】

なお実験では、副写像 $f(m, s)$ を計算する際、もし λ が0.1を越えたら $f(m, s)$ の計算は止めて $f(m, s+1)$ の計算に移行した。 $\lambda > 0.1$ のとき、画素の輝度255レベル中のわずか「3」の違いが副写像の計算に影響したためであり、 $\lambda > 0.1$ のとき正しい結果を得ることは困難だったためである。

【0089】

[1. 4. 2] ヒストグラム $h(1)$

$C(m, s)_f$ の検査はヒストグラム $h(1)$ に依存しない。全単射及びその第3の条件の検査の際、 $h(1)$ に影響を受けうる。実際に $(\lambda, C(m, s)_f)$ をプロットすると、 k は通常1付近にある。実験では $k=1$ を用い、 $B_0 \lambda^2$ と $B_1 \lambda^2$ を検査した。仮に k の本当の値が1未満であれば、 $B_0 \lambda^2$ と $B_1 \lambda^2$ は定数にならず、ファクター $\lambda^{(1-k)/2}$ に従って徐々に増加する。 $h(1)$ が定数であれば、例えばファクターは $\lambda^{1/2}$ である。しかし、こうした差は閾値 $B_{0 \text{ thresh}}$ を正しく設定することによって吸収することができる。

【0090】

ここで次式のごとく始点画像を中心が (x_0, y_0) 、半径 r の円形のオブジェクトであると仮定する。

【数 3 2】

$$p(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_0)^2 + (j-y_0)^2}) & (\sqrt{(i-x_0)^2 + (j-y_0)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 2})$$

一方、終点画像は、次式のごとく中心 (x_1, y_1) 、半径が r のオブジェクトであるとする。

【数 3 3】

$$q(i,j) = \begin{cases} \frac{255}{r} c(\sqrt{(i-x_1)^2 + (j-y_1)^2}) & (\sqrt{(i-x_1)^2 + (j-y_1)^2} \leq r) \\ 0 & (\text{otherwise}) \end{cases} \quad (\text{式 3 3})$$

ここで $c(x)$ は $c(x) = x^k$ の形であるとする。中心 (x_0, y_0) 及び (x_1, y_1) が十分遠い場合、ヒストグラム $h(l)$ は次式の形となる。

【数 3 4】

$$h(l) \propto r l^k \quad (k \neq 0) \quad (\text{式 3 4})$$

$k = 1$ のとき、画像は背景に埋め込まれた鮮明な境界線を持つオブジェクトを示す。このオブジェクトは中心が暗く、周囲にいくに従って明るくなる。 $k = -1$ のとき、画像は曖昧な境界線を持つオブジェクトを表す。このオブジェクトは中心が最も明るく、周囲にいくに従って暗くなる。一般のオブジェクトはこれらふたつのタイプのオブジェクトの間にあると考えてもさして一般性を失わない。したがって、 k は $-1 \leq k \leq 1$ として大抵の場合をカバーでき、式 27 が一般に減少関数であることが保障される。

【0091】

なお、式 34 からわかるように、 r は画像の解像度に影響されること、すなわち r は 2^m に比例することに注意すべきである。このために [1. 4. 1] においてファクター 2^m を導入した。

【0092】

[1. 4. 3] η の動的決定

パラメータ η も同様の方法で自動決定できる。はじめに $\eta = 0$ とし、最も細かい解像度における最終的な写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算する。つづいて、 η をある値 $\Delta \eta$ だけ増加させ、再び最も細かい解像度における最終写像 $f^{(n)}$ 及びエネルギー $C_f^{(n)}$ を計算し直す。この過程を最適値が求まる

まで続ける。 η は写像の剛性を示す。次式の重みだからである。

【0093】

【数35】

$$E_{0(i,j)}^{(m,s)} = \|f^{(m,s)}(i,j) - f^{(m,s-1)}(i,j)\|^2$$

(式35)

η が0のとき、 $D^{(n)}_f$ は直前の副写像と無関係に決定され、現在の副写像は弾性的に変形され、過度に歪むことになる。一方、 η が非常に大きな値のとき、 $D^{(n)}_f$ は直前の副写像によってほぼ完全に決まる。このとき副写像は非常に剛性が高く、画素は同じ場所に射影される。その結果、写像は単位写像になる。 η の値が0から次第に増えるとき、後述のごとく $C^{(n)}_f$ は徐々に減少する。しかし η の値が最適値を越えると、図4に示すとおり、エネルギーは増加し始める。同図のX軸は η 、Y軸は C_f である。

【0094】

この方法で $C^{(n)}_f$ を最小にする最適な η の値を得ることができる。しかし、 λ の場合に比べていろいろな要素が計算に影響する結果、 $C^{(n)}_f$ は小さく揺らぎながら変化する。 λ の場合は、入力が微小量変化するたびに副写像を1回計算しなおすだけだが、 η の場合はすべての副写像が計算しなおされるためである。このため、得られた $C^{(n)}_f$ の値が最小であるかどうかを即座に判断することはできない。最小値の候補が見つければ、さらに細かい区間を設定することによって真の最小値を探す必要がある。

【0095】

[1.5] スーパーサンプリング

画素間の対応関係を決定する際、自由度を増やすために、 $f^{(m,s)}$ の値域を $R \times R$ に拡張することができる (R は実数の集合)。この場合、終点画像の画素の輝度が補間され、非整数点、

【数36】

$$V(q_{f^{(m,s)}(i,j)}^{(m,s)}) \quad (\text{式36})$$

における輝度を持つ $f(m, s)$ が提供される。つまりスーパーサンプリングが行われる。実験では、 $f(m, s)$ は整数及び半整数値をとることが許され、

【数 37】

$$V(q_{(i,j)+(0.5,0.5)}^{(m,s)}) \quad (\text{式 37})$$

は、

【数 38】

$$(V(q_{(i,j)}^{(m,s)}) + V(q_{(i,j)+(1,1)}^{(m,s)}))/2 \quad (\text{式 38})$$

によって与えられた。

【0096】

[1. 6] 各画像の画素の輝度の正規化

始点画像と終点画像がきわめて異なるオブジェクトを含んでいるとき、写像の計算に元の画素の輝度がそのままでは利用しにくい。輝度の差が大きいために輝度に関するエネルギー $C(m, s)_f$ が大きくなりすぎ、正しい評価がしづらいためである。

【0097】

例えば、人の顔と猫の顔のマッチングをとる場合を考える。猫の顔は毛で覆われており、非常に明るい画素と非常に暗い画素が混じっている。この場合、ふたつの顔の間の副写像を計算するために、まず副画像を正規化する。すなわち、最も暗い画素の輝度を 0、最も明るいそれを 255 に設定し、他の画素の輝度は線形補間によって求めておく。

【0098】

[1. 7] インプリメンテーション

始点画像のスキャンに従って計算がリニアに進行する帰納的な方法を用いる。始めに、1 番上の左端の画素 $(i, j) = (0, 0)$ について $f(m, s)$ の値を決定する。次に i を 1 ずつ増やしながらか $f(m, s)(i, j)$ の値を決定する。 i の値が画像の幅に到達したとき、 j の値を 1 増やし、 i を 0 に戻す。以降、始点画像のスキャンに伴い $f(m, s)(i, j)$ を決定していく。すべて

の点について画素の対応が決まれば、ひとつの写像 $f(m, s)$ が決まる。

【0099】

ある $p(i, j)$ について対応点 $q_f(i, j)$ が決まれば、つぎに $p(i, j+1)$ の対応点 $q_f(i, j+1)$ が決められる。この際、 $q_f(i, j+1)$ の位置は全単射条件を満たすために、 $q_f(i, j)$ の位置によって制限される。したがって、先に対応点が決まる点ほどこのシステムでは優先度が高くなる。つねに $(0, 0)$ が最も優先される状態がつづく、求められる最終の写像に余計な偏向が加わる。本前提技術ではこの状態を回避するために、 $f(m, s)$ を以下の方法で決めていく。

【0100】

まず $(s \bmod 4)$ が0の場合、 $(0, 0)$ を開始点とし i 及び j を徐々に増やしながらかめていく。 $(s \bmod 4)$ が1の場合、最上行の右端点を開始点とし、 i を減少、 j を増加させながらかめていく。 $(s \bmod 4)$ が2のとき、最下行の右端点を開始点とし、 i 及び j を減少させながらかめていく。 $(s \bmod 4)$ が3の場合、最下行の左端点を開始点とし、 i を増加、 j を減少させながらかめていく。解像度が最も細かい第 n レベルには副写像という概念、すなわちパラメータ s が存在しないため、仮に $s=0$ 及び $s=2$ であるとしてふたつの方向を連続的に計算した。

【0101】

実際のインプリメンテーションでは、全単射条件を破る候補に対してペナルティを与えることにより、候補 $(k, 1)$ の中からできる限り全単射条件を満たす $f(m, s)(i, j)$ ($m=0, \dots, n$) の値を選んだ。第3の条件を破る候補のエネルギー $D(k, 1)$ には ϕ を掛け、一方、第1または第2の条件を破る候補には ψ を掛ける。今回は $\phi=2$ 、 $\psi=100000$ を用いた。

【0102】

前述の全単射条件のチェックのために、実際の手続として $(k, 1) = f(m, s)(i, j)$ を決定する際に以下のテストを行った。すなわち $f(m, s)(i, j)$ の相続四辺形に含まれる各格子点 $(k, 1)$ に対し、次式の外積の z 成分が0以上になるかどうかを確かめる。

【0103】

【数39】

$$W = \vec{A} \times \vec{B}$$

(式39)

ただしここで、

【数40】

$$\vec{A} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{f(m,s)}^{(m,s)}(i+1,j-1)}$$

(式40)

【数41】

$$\vec{B} = \overrightarrow{q_{f(m,s)}^{(m,s)}(i,j-1) q_{(k,l)}^{(m,s)}}$$

(式41)

である（ここでベクトルは三次元ベクトルとし、z軸は直交右手座標系において定義される）。もしWが負であれば、その候補についてはD(m, s)(k, 1)にφを掛けることによってペナルティを与え、できるかぎり選択しないようにする。

【0104】

図5(a)、図5(b)はこの条件を検査する理由を示している。図5(a)はペナルティのない候補、図5(b)はペナルティがある候補をそれぞれ表す。隣接画素(i, j+1)に対する写像f(m, s)(i, j+1)を決定する際、Wのz成分が負であれば始点画像平面上において全単射条件を満足する画素は存在しない。なぜなら、q(m, s)(k, 1)は隣接する四辺形の境界線を越えるためである。

【0105】

[1. 7. 1] 副写像の順序

インプリメンテーションでは、解像度レベルが偶数のときにはσ(0)=0、σ(1)=1、σ(2)=2、σ(3)=3、σ(4)=0を用い、奇数のときはσ(0)=3、σ(1)=2、σ(2)=1、σ(3)=0、σ(4)=3を用いた。このことで、副写像を適度にシャッフルした。なお、本来副写像は4種類であり、sは0~3のいずれかである。しかし、実際にはs=4に相当する処

理を行った。その理由は後述する。

【0106】

[1. 8] 補間計算

始点画像と終点画像の間の写像が決定された後、対応しあう画素の輝度が補間される。実験では、トライリニア補間を用いた。始点画像平面における正方形 $p(i, j) \ p(i+1, j) \ p(i, j+1) \ p(i+1, j+1)$ が終点画像平面上の四辺形 $q_f(i, j) \ q_f(i+1, j) \ q_f(i, j+1) \ q_f(i+1, j+1)$ に射影されると仮定する。簡単のため、画像間の距離を1とする。始点画像平面からの距離が t ($0 \leq t \leq 1$) である中間画像の画素 $r(x, y, t)$ ($0 \leq x \leq N-1, 0 \leq y \leq M-1$) は以下の要領で求められる。まず画素 $r(x, y, t)$ の位置 (ただし $x, y, t \in R$) を次式で求める。

【0107】

【数42】

$$\begin{aligned} (x, y) = & (1-dx)(1-dy)(1-t)(i, j) + (1-dx)(1-dy)tf(i, j) \\ & + dx(1-dy)(1-t)(i+1, j) + dx(1-dy)tf(i+1, j) \\ & + (1-dx)dy(1-t)(i, j+1) + (1-dx)dytf(i, j+1) \\ & + dx dy(1-t)(i+1, j+1) + dx dy tf(i+1, j+1) \end{aligned}$$

(式42)

つづいて $r(x, y, t)$ における画素の輝度が次の式を用いて決定される。

【0108】

【数43】

$$\begin{aligned} V(r(x, y, t)) = & (1-dx)(1-dy)(1-t)V(p_{(i,j)}) + (1-dx)(1-dy)tV(q_{f(i,j)}) \\ & + dx(1-dy)(1-t)V(p_{(i+1,j)}) + dx(1-dy)tV(q_{f(i+1,j)}) \\ & + (1-dx)dy(1-t)V(p_{(i,j+1)}) + (1-dx)dytV(q_{f(i,j+1)}) \\ & + dx dy(1-t)V(p_{(i+1,j+1)}) + dx dy tV(q_{f(i+1,j+1)}) \end{aligned}$$

(式43)

ここで dx 及び dy はパラメータであり、0から1まで変化する。

【0109】

[1. 9] 拘束条件を課したときの写像

いままでは拘束条件がいっさい存在しない場合の写像の決定を述べた。しかし、始点画像と終点画像の特定の画素間に予め対応関係が規定されているとき、これを拘束条件としたうえで写像を決定することができる。

【0110】

基本的な考えは、まず始点画像の特定の画素を終点画像の特定の画素に移す大まかな写像によって始点画像を大まかに変形し、しかる後、写像 f を正確に計算する。

【0111】

まず始めに、始点画像の特定の画素を終点画像の特定の画素に射影し、始点画像の他の画素を適当な位置に射影する大まかな写像を決める。すなわち、特定の画素に近い画素は、その特定の画素が射影される場所の近くに射影されるような写像である。ここで第 m レベルの大まかな写像を $F^{(m)}$ と記述する。

【0112】

大まかな写像 F は以下の要領で決める。まず、いくつかの画素について写像を特定する。始点画像について n_s 個の画素、

【数44】

$$p(i_0, j_0), p(i_1, j_1), \dots, p(i_{n_s-1}, j_{n_s-1})$$

(式44)

を特定するとき、以下の値を決める。

【数45】

$$\begin{aligned} F^{(n)}(i_0, j_0) &= (k_0, l_0), \\ F^{(n)}(i_1, j_0) &= (k_1, l_1), \dots, \end{aligned}$$

$$F^{(n)}(i_{n_s-1}, j_{n_s-1}) = (k_{n_s-1}, l_{n_s-1})$$

(式45)

始点画像の他の画素の変位量は、 $p(i_h, j_h)$ ($h = 0, \dots, n_s - 1$) の変位に重み付けをして求められる平均である。すなわち画素 $p(i, j)$ は、終点画像の以下の画素に射影される。

【0113】

【数 4 6】

$$F^{(m)}(i, j) = \frac{(i, j) + \sum_{h=0}^{h=n_s-1} (k_h - i_h, l_h - j_h) \text{weight}_h(i, j)}{2^{n-m}} \quad (\text{式 4 6})$$

ただしここで、

【数 4 7】

$$\text{weight}_h(i, j) = \frac{1/\|(i_h - i, j_h - j)\|^2}{\text{total weight}(i, j)} \quad (\text{式 4 7})$$

【数 4 8】

$$\text{total weight}(i, j) = \sum_{h=0}^{h=n_s-1} 1/\|(i_h - i, j_h - j)\|^2 \quad (\text{式 4 8})$$

とする。

【0114】

つづいて、 $F^{(m)}$ に近い候補写像 f がより少ないエネルギーを持つように、その写像 f のエネルギー $D^{(m, s)}(i, j)$ を変更する。正確には、 $D^{(m, s)}(i, j)$ は、

【数 4 9】

$$D_{(i, j)}^{(m, s)} = E_{0(i, j)}^{(m, s)} + \eta E_{1(i, j)}^{(m, s)} + \kappa E_{2(i, j)}^{(m, s)} \quad (\text{式 4 9})$$

である。ただし、

【数 5 0】

$$E_{2(i, j)}^{(m, s)} = \begin{cases} 0, & \text{if } \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2 \leq \lfloor \frac{\rho^2}{2^{2(n-m)}} \rfloor \\ \|F^{(m)}(i, j) - f^{(m, s)}(i, j)\|^2, & \text{otherwise} \end{cases} \quad (\text{式 5 0})$$

であり、 $\kappa, \rho \geq 0$ とする。最後に、前述の写像の自動計算プロセスにより、 f を完全に決定する。

【0115】

ここで、 $f^{(m, s)}(i, j)$ が $F^{(m)}(i, j)$ に十分近いとき、つまりそれ

らの距離が、

【数 5 1】

$$\left[\frac{\rho^2}{2^{2(n-m)}} \right] \quad (\text{式 5 1})$$

以内であるとき、 $E_2(m, s)(i, j)$ が 0 になることに注意すべきである。そのように定義した理由は、各 $f(m, s)(i, j)$ が $F(m)(i, j)$ に十分近い限り、終点画像において適切な位置に落ち着くよう、その値を自動的に決めたいためである。この理由により、正確な対応関係を詳細に特定する必要がなく、始点画像は終点画像にマッチするように自動的にマッピングされる。

【0 1 1 6】

[2] 具体的な処理手順

[1] の各要素技術による処理の流れを説明する。

図 6 は前提技術の全体手順を示すフローチャートである。同図のごとく、まず多重解像度特異点フィルタを用いた処理を行い (S 1)、つづいて始点画像と終点画像のマッチングをとる (S 2)。ただし、S 2 は必須ではなく、S 1 で得られた画像の特徴をもとに画像認識などの処理を行ってもよい。

【0 1 1 7】

図 7 は図 6 の S 1 の詳細を示すフローチャートである。ここでは S 2 で始点画像と終点画像のマッチングをとることを前提としている。そのため、まず特異点フィルタによって始点画像の階層化を行い (S 1 0)、一連の始点階層画像を得る。つづいて同様の方法で終点画像の階層化を行い (S 1 1)、一連の終点階層画像を得る。ただし、S 1 0 と S 1 1 の順序は任意であるし、始点階層画像と終点階層画像を並行して生成していくこともできる。

【0 1 1 8】

図 8 は図 7 の S 1 0 の詳細を示すフローチャートである。もとの始点画像のサイズは $2^n \times 2^n$ とする。始点階層画像は解像度が細かいほうから順に作られるため、処理の対象となる解像度レベルを示すパラメータ m を n にセットする (S 1 0 0)。つづいて第 m レベルの画像 $p(m, 0)$ 、 $p(m, 1)$ 、 $p(m, 2)$ 、 $p(m, 3)$ から特異点フィルタを用いて特異点を検出し (S 1 0 1)、そ

れぞれ第 $m-1$ レベルの画像 $p(m-1, 0)$ 、 $p(m-1, 1)$ 、 $p(m-1, 2)$ 、 $p(m-1, 3)$ を生成する (S102)。ここでは $m=n$ であるため、 $p(m, 0) = p(m, 1) = p(m, 2) = p(m, 3) = p(n)$ であり、ひとつの始点画像から4種類の副画像が生成される。

【0119】

図9は第 m レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示している。同図の数値は各画素の輝度を示す。同図の $p(m, s)$ は $p(m, 0) \sim p(m, 3)$ の4つの画像を象徴するもので、 $p(m-1, 0)$ を生成する場合には、 $p(m, s)$ は $p(m, 0)$ であると考え。[1. 2]で示した規則により、 $p(m-1, 0)$ は例えば同図で輝度を記入したブロックについて、そこに含まれる4画素のうち「3」、 $p(m-1, 1)$ は「8」、 $p(m-1, 2)$ は「6」、 $p(m-1, 3)$ を「10」をそれぞれ取得し、このブロックをそれぞれ取得したひとつの画素で置き換える。したがって、第 $m-1$ レベルの副画像のサイズは $2^{m-1} \times 2^{m-1}$ になる。

【0120】

つづいて m をデクリメントし (図8のS103)、 m が負になっていないことを確認し (S104)、S101に戻ってつぎに解像度の粗い副画像を生成していく。この繰り返し処理の結果、 $m=0$ 、すなわち第0レベルの副画像が生成された時点でS10が終了する。第0レベルの副画像のサイズは 1×1 である。

【0121】

図10はS10によって生成された始点階層画像を $n=3$ の場合について例示している。最初の始点画像のみが4つの系列に共通であり、以降特異点の種類に応じてそれぞれ独立に副画像が生成されていく。なお、図8の処理は図7のS11にも共通であり、同様の手順を経て終点階層画像も生成される。以上で図6のS1による処理が完了する。

【0122】

前提技術では、図6のS2に進むためにマッチング評価の準備をする。図11はその手順を示している。同図のごとく、まず複数の評価式が設定される (S30)。[1. 3. 2. 1]で導入した画素に関するエネルギー $C(m, s)_f$ と

[1. 3. 2. 2] で導入した写像の滑らかさに関するエネルギー $D(m, s)_f$ がそれである。つぎに、これらの評価式を統合して総合評価式を立てる (S 3 1)。[1. 3. 2. 3] で導入した総エネルギー $\lambda C(m, s)_f + D(m, s)_f$ がそれであり、[1. 3. 2. 2] で導入した η を用いれば、

$$\sum \sum (\lambda C(m, s)(i, j) + \eta E_0(m, s)(i, j) + E_1(m, s)(i, j)) \quad (\text{式52})$$

となる。ただし、総和は i, j についてそれぞれ $0, 1, \dots, 2^m - 1$ で計算する。以上でマッチング評価の準備が整う。

【0123】

図12は図6のS2の詳細を示すフローチャートである。[1]で述べたごとく、始点階層画像と終点階層画像のマッチングは互いに同じ解像度レベルの画像どうしでとられる。画像間の大域的なマッチングを良好にとるために、解像度が粗いレベルから順にマッチングを計算する。特異点フィルタを用いて始点階層画像および終点階層画像を生成しているため、特異点の位置や輝度は解像度の粗いレベルでも明確に保存されており、大域的なマッチングの結果は従来に比べて非常に優れたものになる。

【0124】

図12のごとく、まず係数パラメータ η を0、レベルパラメータ m を0に設定する (S 2 0)。つづいて、始点階層画像中の第 m レベルの4つの副画像と終点階層画像中の第 m レベルの4つの副画像のそれぞれの間でマッチングを計算し、それぞれ全単射条件を満たし、かつエネルギーを最小にするような4種類の副写像 $f(m, s)$ ($s = 0, 1, 2, 3$) を求める (S 2 1)。全単射条件は[1. 3. 3]で述べた相続四辺形を用いて検査される。この際、式17、18が示すように、第 m レベルにおける副写像は第 $m-1$ レベルのそれらに拘束されるため、より解像度の粗いレベルにおけるマッチングが順次利用されていく。これは異なるレベル間の垂直的参照である。なお、いま $m=0$ であってそれより粗いレベルはないが、この例外的な処理は図13で後述する。

【0125】

一方、同一レベル内における水平的参照も行われる。[1. 3. 3]の式20

のごとく、 $f(m, 3)$ は $f(m, 2)$ に、 $f(m, 2)$ は $f(m, 1)$ に、 $f(m, 1)$ は $f(m, 0)$ に、それぞれ類似するように決める。その理由は、特異点の種類が違って、それらがもともとと同じ始点画像と終点画像に含まれている以上、副写像がまったく異なるという状況は不自然だからである。式20からわかるように、副写像どうしが近いほどエネルギーは小さくなり、マッチングが良好とみなされる。

【0126】

なお、最初に決めるべき $f(m, 0)$ については同一のレベルで参照できる副写像がないため、式19に示すごとくひとつ粗いレベルを参照する。ただし、実験では $f(m, 3)$ まで求まった後、これを拘束条件として $f(m, 0)$ を一回更新するという手続をとった。これは式20に $s=4$ を代入し、 $f(m, 4)$ を新たな $f(m, 0)$ とすることに等しい。 $f(m, 0)$ と $f(m, 3)$ の関連度が低くなり過ぎる傾向を回避するためであり、この措置によって実験結果がより良好になった。この措置に加え、実験では[1. 7. 1]に示す副写像のシャッフルも行った。これも本来特異点の種類ごとに決まる副写像どうしの関連度を密接に保つ趣旨である。また、処理の開始点に依存する偏向を回避するために、 s の値にしたがって開始点の位置を変える点は[1. 7]で述べたとおりである。

【0127】

図13は第0レベルにおいて副写像を決定する様子を示す図である。第0レベルでは各副画像がただひとつの画素で構成されるため、4つの副写像 $f(0, s)$ はすべて自動的に単位写像に決まる。図14は第1レベルにおいて副写像を決定する様子を示す図である。第1レベルでは副画像がそれぞれ4画素で構成される。同図ではこれら4画素が実線で示されている。いま、 $p(1, s)$ の点 x の対応点を $q(1, s)$ の中で探すとき、以下の手順を踏む。

【0128】

1. 第1レベルの解像度で点 x の左上点 a 、右上点 b 、左下点 c 、右下点 d を求める。

2. 点 $a \sim d$ がひとつ粗いレベル、つまり第0レベルにおいて属する画素を探す。図14の場合、点 $a \sim d$ はそれぞれ画素 $A \sim D$ に属する。ただし、画素 $A \sim$

Cは本来存在しない仮想的な画素である。

3. 第0レベルですでに求まっている画素A～Dの対応点A'～D'を $q(1, s)$ の中にプロットする。画素A'～C'は仮想的な画素であり、それぞれ画素A～Cと同じ位置にあるものとする。

4. 画素Aの中の点aの対応点a'が画素A'の中にあるとみなし、点a'をプロットする。このとき、点aが画素Aの中で占める位置（この場合、右下）と、点a'が画素A'の中で占める位置が同じであると仮定する。

5. 4と同様の方法で対応点b'～d'をプロットし、点a'～d'で相続四辺形を作る。

6. 相続四辺形の中でエネルギーが最小になるよう、点xの対応点x'を探す。対応点x'の候補として、例えば画素の中心が相続四辺形に含まれるものに限定してもよい。図14の場合、4つの画素がすべて候補になる。

【0129】

以上がある点xの対応点の決定手順である。同様の処理を他のすべての点について行い、副写像を決める。第2レベル以上のレベルでは、次第に相続四辺形の形が崩れていくと考えられるため、図3に示すように画素A'～D'の間隔が空いていく状況が発生する。

【0130】

こうして、ある第mレベルの4つの副写像が決まれば、mをインクリメントし（図12のS22）、mがnを超えていないことを確かめて（S23）、S21に戻る。以下、S21に戻るたびに次第に細かい解像度のレベルの副写像を求め、最後にS21に戻ったときに第nレベルの写像 $f^{(n)}$ を決める。この写像は $\eta=0$ に関して定まったものであるから、 $f^{(n)}(\eta=0)$ と書く。

【0131】

つぎに異なる η に関する写像も求めるべく、 η を $\Delta\eta$ だけシフトし、mをゼロクリアする（S24）。新たな η が所定の探索打ち切り値 η_{max} を超えていないことを確認し（S25）、S21に戻り、今回の η に関して写像 $f^{(n)}(\eta=\Delta\eta)$ を求める。この処理を繰り返し、S21で $f^{(n)}(\eta=i\Delta\eta)$ （ $i=0, 1, \dots$ ）を求めていく。 η が η_{max} を超えたときS26に進み、後述の方

法で最適な $\eta = \eta_{opt}$ を決定し、 $f^{(n)}$ ($\eta = \eta_{opt}$) を最終的に写像 $f^{(n)}$ とする。

【0132】

図15は図12のS21の詳細を示すフローチャートである。このフローチャートにより、ある定まった η について、第 m レベルにおける副写像が決まる。副写像を決める際、前提技術では副写像ごとに最適な λ を独立して決める。

【0133】

同図のごとく、まず s と λ をゼロクリアする (S210)。つぎに、そのときの λ について (および暗に η について) エネルギーを最小にする副写像 $f^{(m, s)}$ を求め (S211)、これを $f^{(m, s)}$ ($\lambda = 0$) と書く。異なる λ に関する写像も求めるべく、 λ を $\Delta\lambda$ だけシフトし、新たな λ が所定の探索打切り値 λ_{max} を超えていないことを確認し (S213)、S211に戻り、以降の繰り返し処理で $f^{(m, s)}$ ($\lambda = i\Delta\lambda$) ($i = 0, 1, \dots$) を求める。 λ が λ_{max} を超えたときS214に進み、最適な $\lambda = \lambda_{opt}$ を決定し、 $f^{(m, s)}$ ($\lambda = \lambda_{opt}$) を最終的に写像 $f^{(m, s)}$ とする (S214)。

【0134】

つぎに、同一レベルにおける他の副写像を求めるべく、 λ をゼロクリアし、 s をインクリメントする (S215)。 s が4を超えていないことを確認し (S216)、S211に戻る。 $s = 4$ になれば上述のごとく $f^{(m, 3)}$ を利用して $f^{(m, 0)}$ を更新し、そのレベルにおける副写像の決定を終了する。

【0135】

図16は、ある m と s について λ を変えながら求められた $f^{(m, s)}$ ($\lambda = i\Delta\lambda$) ($i = 0, 1, \dots$) に対応するエネルギー $C^{(m, s)}_f$ の挙動を示す図である。[1. 4]で述べたとおり、 λ が増加すると通常 $C^{(m, s)}_f$ は減少する。しかし、 λ が最適値を超えると $C^{(m, s)}_f$ は増加に転じる。そこで本前提技術では $C^{(m, s)}_f$ が極小値をとるときの λ を λ_{opt} と決める。同図のように $\lambda > \lambda_{opt}$ の範囲で再度 $C^{(m, s)}_f$ が小さくなっていても、その時点ではすでに写像がくずれていて意味をなさないため、最初の極小点に注目すればよい。 λ_{opt} は副写像ごとに独立して決めていき、最後に $f^{(n)}$ に

についてもひとつ定まる。

【0136】

一方、図17は、 η を変えながら求められた $f^{(n)}$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギー $C^{(n)}_f$ の挙動を示す図である。ここでも η が増加すると通常 $C^{(n)}_f$ は減少するが、 η が最適値を超えると $C^{(n)}_f$ は増加に転じる。そこで $C^{(n)}_f$ が極小値をとるときの η を η_{opt} と決める。図17は図4の横軸のゼロ付近を拡大した図と考えてよい。 η_{opt} が決まれば $f^{(n)}$ を最終決定することができる。

【0137】

以上、本前提技術によれば種々のメリットが得られる。まずエッジを検出する必要がないため、エッジ検出タイプの従来技術の課題を解消できる。また、画像に含まれるオブジェクトに対する先験的な知識も不要であり、対応点の自動検出が実現する。特異点フィルタによれば、解像度の粗いレベルでも特異点の輝度や位置を維持することができ、オブジェクト認識、特徴抽出、画像マッチングに極めて有利である。その結果、人手作業を大幅に軽減する画像処理システムの構築が可能となる。

【0138】

なお、本前提技術について次のような変形技術も考えられる。

(1) 前提技術では始点階層画像と終点階層画像の間でマッチングをとる際にパラメータの自動決定を行ったが、この方法は階層画像間ではなく、通常の2枚の画像間のマッチングをとる場合全般に利用できる。

【0139】

たとえば2枚の画像間で、画素の輝度の差に関するエネルギー E_0 と画素の位置的なずれに関するエネルギー E_1 のふたつを評価式とし、これらの線形和 $E_{tot} = \alpha E_0 + E_1$ を総合評価式とする。この総合評価式の極値付近に注目して α を自動決定する。つまり、いろいろな α について E_{tot} が最小になるような写像を求める。それらの写像のうち、 α に関して E_1 が極小値をとるときの α を最適パラメータと決める。そのパラメータに対応する写像を最終的に両画像間の最適マッチングとみなす。

【0140】

これ以外にも評価式の設定にはいろいろな方法があり、例えば $1/E_1$ と $1/E_2$ のように、評価結果が良好なほど大きな値をとるものを採用してもよい。総合評価式も必ずしも線形和である必要はなく、 n 乗和 ($n=2$ 、 $1/2$ 、 -1 、 -2 など)、多項式、任意の関数などを適宜選択すればよい。

【0141】

パラメータも、 α のみ、前提技術のごとく η と λ のふたつの場合、それ以上の場合など、いずれでもよい。パラメータが3以上の場合はひとつずつ変化させて決めていく。

【0142】

(2) 本前提技術では、総合評価式の値が最小になるよう写像を決めた後、総合評価式を構成するひとつの評価式である $C(m, s)_f$ が極小になる点を検出してパラメータを決定した。しかし、こうした二段回処理の代わりに、状況によっては単に総合評価式の最小値が最小になるようにパラメータを決めても効果的である。その場合、例えば $\alpha E_0 + \beta E_1$ を総合評価式とし、 $\alpha + \beta = 1$ なる拘束条件を設けて各評価式を平等に扱うなどの措置を講じてもよい。パラメータの自動決定の本質は、エネルギーが最小になるようにパラメータを決めていく点にあるからである。

【0143】

(3) 前提技術では各解像度レベルで4種類の特異点に関する4種類の副画像を生成した。しかし、当然4種類のうち1、2、3種類を選択的に用いてもよい。例えば、画像中に明るい点がひとつだけ存在する状態であれば、極大点に関する $f(m, 3)$ だけで階層画像を生成しても相応の効果が得られるはずである。その場合、同一レベルで異なる副写像は不要になるため、 s に関する計算量が減る効果がある。

【0144】

(4) 本前提技術では特異点フィルタによってレベルがひとつ進むと画素が $1/4$ になった。例えば 3×3 で1ブロックとし、その中で特異点を探す構成も可能であり、その場合、レベルがひとつ進むと画素は $1/9$ になる。

【0145】

(5) 始点画像と終点画像がカラーの場合、それらをまず白黒画像に変換し、写像を計算する。その結果求められた写像を用いて始点のカラー画像を変換する。それ以外の方法として、RGBの各成分について副写像を計算してもよい。

【0146】

[映像効果または画像補間に関する実施の形態]

以上の前提技術を利用した実施の形態を説明する。実施の形態1はより原理に近いもの、実施の形態2以降はよりアプリケーションに近いものを挙げるが、それらに明確な境界線を引く必要はなく、いずれかの実施の形態の任意の機能と他の実施の形態の任意の機能が組み合わされてもよい。

【0147】

実施の形態1

図18は、キーフレームである第1画像I1、第2画像I2を示し、それらの間で、ある画素 $p_1(x_1, y_1)$ と $p_2(x_2, y_2)$ が対応している。これらの対応は前提技術で求められている。

【0148】

図19は、第1画像I1上にメッシュを切り、そのメッシュを構成するポリゴンの第2画像I2上の対応位置を示す。いま第1画像I1上で注目するポリゴンR1は4個の格子点A、B、C、Dによって決まっている。このポリゴンR1を「起点ポリゴン」と名付ける。これら4個の格子点A、B、C、Dは、図18に示したごとく、それぞれ第2画像I2上に対応する点A'、B'、C'、D'を有し、これらの点で形成されるポリゴンR2を「終点ポリゴン」と名付ける。起点ポリゴンは一般に長方形であるが、終点ポリゴンは一般に四辺形である。いずれにせよ、この実施の形態では、画素単位で第1画像I1と第2画像I2の対応関係を記述せず、起点ポリゴンの格子点について、その対応画素を記述する。その記述は対応点ファイルに書き込まれる。格子点に注目することにより、対応点ファイルの容量を非常に小さくすることができる。

【0149】

対応点ファイルは第1画像I1と第2画像I2の中間画像を生成するために利

用される。この点は前提技術で述べたごとく、対応点どうしの位置を補間することにより、任意の時間的位置の中間画像を生成できる。したがって、第1画像I1、第2画像I2、および対応点ファイルを保存しておくことで、それに2つの画像間のモーフィングや滑らかな動画を生成することができる。このため、動画の圧縮効果が得られる。

【0150】

図20は対応点ファイルから格子点以外の点に関する対応関係を計算する方法を示す。対応点ファイルには格子点に関する情報しかないので、ポリゴン内部の点について対応情報を別途計算する必要がある。図20では、図19の起点ポリゴンR1の下半分に当たる三角形ABCと、同様に終点ポリゴンR2の下半分に当たる三角形A'B'C'との対応を示す。いま、起点ポリゴン側の三角形ABC内部の点Qが、線分ACを $t : (1-t)$ に内分し、その内分点と点Bの間を $s : (1-s)$ に内分する位置にあったとする。この点Qの終点ポリゴン側の三角形A'B'C'における対応点Q'は、線分A'C'を $t : (1-t)$ に内分し、その内分点と点B'の間を $s : (1-s)$ に内分する位置にあると考えればよい。すなわち、起点ポリゴンを三角形に分割し、その三角形に関するベクトルの内分の形式にて終点ポリゴン内部の点を求めればよい。ベクトル斜体で示せば

$$\vec{BQ} = (1-s) \{ (1-t) \vec{BA} + t \vec{BC} \}$$

であり、したがって、

$$\vec{B'Q'} = (1-s) \{ (1-t) \vec{B'A'} + t \vec{B'C'} \}$$

当然ながら、同様の処理を起点ポリゴンR1の上半分に当たる三角形ACDと、同様に終点ポリゴンR2の上半分に当たる三角形A'C'D'との間についてもなすとする。

【0151】

図21は、以上の処理手順を示す。まず図19のごとく、第1画像I1上に取りられた格子点に関するマッチング結果を取得する(S10)。このとき、前提技術の画素単位のマッチングを実行し、その結果の中から格子点に当たる部分を抽出すればよい。ただし、前提技術を用いずに他のマッチング技術、例えばオブテ

イカルフローやブロックマッチングをもとに格子点に関するマッチング結果を特定してもよい。

【0152】

つづいて、図19の右側のごとく、第2画像I2上に終点ポリゴンを定義する(S12)。以上で対応点ファイルが生成できるので、このファイルに第1画像I1、第2画像I2を識別するための情報を盛り込んで出力する(S14)。ふたつの画像および対応点ファイルは、任意の記録装置または媒体に保持され、または直接ネットワークや放送波に乗せて伝送されればよい。

【0153】

図22は、対応点ファイルを用いて中間画像を生成する手順を示す。まず、第1画像I1、第2画像I2が読み込まれ(S20)、対応点ファイルが読み込まれる(S22)。つづいて、起点ポリゴン内の点と終点ポリゴン内の点の対応関係が図20の方法によって算出される(S24)。この時点で、画像内のすべての画素に関する対応関係を取得することができる。したがって、前提技術で述べたごとく、互いに対応し合う点の座標と色を $u : (1 - u)$ で内分すれば、第1画像I1と第2画像I2の間を時間的に $(1 - u) : u$ で内分する位置における中間画像を生成することができる(S26)。なお、内挿補間だけでなく、外挿補間をしてもよい。

【0154】

図23は以上の処理を行う映像効果装置10の構成を示す。この装置10は、外部の記憶装置や撮影カメラ等から第1画像I1、第2画像I2を取得する画像入力部12と、それらの画像に前提技術その他の技術によってマッチング計算を施すマッチングプロセッサ14と、それによって生成された対応点ファイルFを格納する対応点ファイル保持部16と、第1画像I1および第2画像I2から中間画像を生成する中間画像生成部18と、第1画像I1、第2画像I2および中間画像をタイミング調整しながらもとの動画に近い映像として表示する表示部20を含む。また、通信部22は、外部からの要望にしたがい、第1画像I1、第2画像I2および対応点ファイルFをネットワークその他の伝送インフラへ送出する。なお、マッチングプロセッサ14には、メッシュの大きさや格子点の位置

などを示すメッシュ情報が入力されている。

【0155】

以上の構成により、画像入力部12に入力された第1画像I1と第2画像I2はマッチングプロセッサ14へ送られる。マッチングプロセッサ14は、それらの画像間で画素単位のマッチング計算を行う。マッチングプロセッサ14はメッシュ情報をもとに対応点ファイルFを生成し、対応点ファイル保持部16へ出力する。

【0156】

中間画像生成部18は、ユーザの求めに応じ、またはその他の要因で対応点ファイルFを読み出し、中間画像を生成する。この中間画像は表示部20へ送られ、そこで画像出力の時間調整が行われ、動画またはモーフィング画像が表示される。この動作からもわかるとおり、中間画像生成部18と表示部20は、本装置10とは別構成のリモート端末側にあってもよく、その場合、その端末が第1画像I1、第2画像I2、対応点ファイルFという比較的軽いデータを受け、自主的に動画を再生できる。

【0157】

通信部22はそのための構成であり、ここではすでにリモート端末がネットワークの先で待っていることを想定している。すなわち、通信部22は第1画像I1、第2画像I2、対応点ファイルFをネットワークや放送波に乗せて送り出し、リモート端末側で動画の表示が行われる。もちろんリモート端末は、表示ではなく、ストレージを目的としてもよい。なお本装置は、ネットワーク等の外部から通信部22を介して第1画像I1、第2画像I2およびそれらの対応点ファイルを入力し、これらを中間画像生成部18に渡して補間計算をし、中間画像を生成する用途もある。図23にはそのためのデータの経路Pが記述されている。

【0158】

実施の形態1の処理内容について実験を行ったところ、第1、第2画像として、例えば256×256程度の大きさを採用し、格子点を縦横10～数十画素おき程度に設定することにより、きわめて良好なモーフィングまたは動画圧縮効果が得られた。対応点ファイルの大きさは数キロから10キロバイト程度であり、

画質の高さとデータ量の小ささが実現できることが確認できた。

【0159】

実施の形態 2

実施の形態 1 で実現された映像効果装置は、電子商取引における商品プレゼンテーションに利用できる。実施の形態 2 のこの装置（以下「商品提示装置」という）は、いくつかの角度から撮影された商品の写真のうち近接しあう 2 枚の画像を第 1 画像 I 1、第 2 画像 I 2 とし、それらの間を滑らかに補間することで商品をユーザに示す。たとえば、10° おきに商品のまわりから撮影した写真を順次補間することで 360° 滑らかな回転画像が得られる。従来、商品の滑らかな回転画像はせいぜい 1、2° ごとに撮影が必要とされたが、実施の形態 1 の技術を利用することで、この角度を 10～20° にできることが判明している。

【0160】

商品提示装置は、実施の形態 1 と異なり、予め生成された対応点ファイル F を利用する。商品の写真は提示前に準備できるためである。図 23 の構成との相違点のみ示せば、画像入力部 12 は第 1 画像 I 1、第 2 画像 I 2 のほかに、対応点ファイル F を取得する。これらは中間画像生成部 18 へ入力され、中間画像が生成される。中間画像は表示部 20 で表示されたり、通信部 22 を介してネットワークその他の伝送インフラへ送出されてもよい。この実施の形態ではマッチングプロセッサ 14 と対応点ファイル保持部 16 は不要である。なお、そうした構成は他の実施の形態にも有効である。

【0161】

商品提示装置は、実施の形態 1 に加え、電子商取引における商品プレゼンテーション機能を提供する。この機能が前述の商品回転機能である場合、ユーザから GUI を通じて入力された指示にしたがって商品を回転させてもよい。そうした指示の方法は、たとえば CG の分野でいろいろと知られており、ここではそうした技術を利用する。

【0162】

機能の別の例として、商品画像を二方向に補間して得られた中間画像を表示してもよい。たとえば、商品を横方向にとびとびに撮影することにくわえ、縦方向

にもとびとびに撮影することにより、横方向だけでなく、縦方向にも補間することができる。この機能は例えば下の工程を含む。

(1) 第1画像、第2画像からなる画像対Aと、それら2枚の画像間の対応点情報Aを取得する工程

(2) 第3画像、第4画像からなる画像対Bと、それら2枚の画像間の対応点情報Bを取得する工程

(3) 画像対Aの2枚の画像間に時間的または空間的に定まる軸Aと画像対Bの2枚の画像間に時間的または空間的に定まる軸Bの位置関係、対応点情報A、および対応点情報Bを利用して、中間画像を補間によって生成する工程。

【0163】

(3)において、軸Aと軸Bを利用して2軸補間 (bilinear interpolation) を行ってもよい。一例として、2つの視点 $p_1(0, 0)$ 、 $p_2(0, 100)$ から得た画像対Aとし、別の2つの視点 $p_3(100, 0)$ 、 $p_4(100, 100)$ から得た画像対Bとする。点 p_1 と p_2 を結ぶ直線が軸A、点 p_3 と p_4 を結ぶ直線が軸Bに当たる。

【0164】

中間画像として、視点 $p' = (50, 50)$ から見た画像を得る場合、まず画像対Aの間の対応点情報Aをもとに視点 $(0, 50)$ からのフレームを生成する。つぎに画像対Bの間の対応点情報Bをもとに視点 $(100, 50)$ からの別のフレームを生成する。つづいて、これらふたつのフレームに補間を施し、すなわちこの場合1:1で内分し、所望の中間フレームが生成される。ここで、縦横両方向の補間をするためには、一般には、軸Aと軸Bが同一直線上にこないよう画像対Aと画像対Bを定めればよい。

【0165】

なお、この例では軸A、軸Bは、それぞれ2枚の画像間に空間的に定まるが、時間的に定まる別の例もある。たとえば、ひとつの視点Pから時刻 $t = t_0$ と t_1 に得たふたつの画像を画像対Aとし、別の視点Qから時刻 $t = t_0$ と t_1 に得たふたつの画像を画像対Bとする。この場合、画像対Aにおいて (P, t_0) で定義される点と (P, t_1) で定義される点を結ぶ直線が軸Aとなり、同様に画

像対Bにおいて (Q, t_0) で定義される点と (Q, t_1) で定義される点を結ぶ直線が軸Bとなる。したがって、中間画像を例えば点 $((P+Q)/2, (t_0+t_1)/2)$ からの画像とすれば、それぞれふたつの軸の中間的な画像を生成した後、これらを補間すればよい。

【0166】

実施の形態3

中間画像の生成に当たり、映像効果に関する機能を提供してもよい。実施の形態1では一般論としての映像効果に触れたが、ここでいう映像効果は、より具体的またはプロフェッショナルに近い用途と考えてもよい。以下機能の例である。

【0167】

3以上の画像の中から隣接するふたつの画像を第1画像および第2画像として選定し、それらの中間画像を生成および表示する。つづいて、別の隣接する画像を第1画像と第2画像とし、以降順に中間画像を生成する。例えば、視点が少しずつ異なる画像S、T、Uがあるとき、まず画像SとTの中間画像を作り、つぎに画像TとUの中間画像を作ることにより、一連の画像を連続的に表示することができる。すなわち、多数の離散的な視点画像間をつぎつぎ補間して中間画像を生成すれば、あたかもパノラマ的な画像が得られる。

【0168】

これと同様の方法であるが、時間の要素を取り入れることもできる。例えば、複数の離散的な画像を、視点を時間とともに動かした画像と捉え、かつそれらの中間画像を順を追って表示すれば、いわゆるウォークスルー画像が得られる。

【0169】

別の機能は、処理の対象となる画像を二方向に補間して得られた中間画像を表示するものである。この方法は実施の形態2でも述べたとおり、例えば二軸補間で実現する。これにより、例えば二方向に連続変化するパノラマ画像も得られる。

【0170】

さらに別の機能は、ある対象物を少しずつ廻りながら撮影し、それらの連続画像を生成する機能である。実施の形態2では商品自体を回転させたが、映像効果

としては、むしろ対象物をぐるりと撮影した画像を表示する点におもしろさがある。一部の映画では、空中にとどまった人を一周する映像がSFX技術で生成されているが、実施の形態1によれば、より少ない写真画像で容易にそうした画像を生成することができる。

【0171】

これら以外にも既存の映像効果を実施の形態に係る技術で効率化し、または高画質化することができる。本実施の形態に係る装置は、いずれの映像効果を行うかをユーザが指定するGUIを搭載し、指定内容に相応しい前処理または後処理の機能を実装しておいてもよい。

【0172】

たとえば、ウォークスルー画像を生成するとき、予め視点の移動経路を指定してもよく、それにしたがって順に中間画像を表示してもよい。二方向に補間する場合は、実施の形態2同様、GUI、例えばマウスの移動方向にしたがって移動する視点を仮定し、補間画像を生成してもよい。マウスが右下から左上へ移動すれば、視点も右下から左上へ移動したと想定し、二軸補間をすればよい。

【0173】

実施の形態4

実施の形態1でも触れたごとく、本発明は動画の圧縮にも向く。少ない静止画像をもとに動画が再生できるため、インターネット等で動画を配信する際に有利である。このため、中間画像の生成に当たり、動画画像の生成に関するユーティリティ機能を提供してもよい。

【0174】

その機能の例として、生成すべき中間画像の枚数を指定できてもよい。たとえば、第1画像と第2画像がもともと一秒の間隔をもって撮影されたものであれば、通常はそれらの間に30枚程度のフレームがある。ユーザがGUI等を通して「中間画像30枚」と指定すれば、時間スケールとしては、もとの動画と同じものが得られる。ここで「3000枚」を指定すれば、時間を100倍に引き延ばしたスローモーション画像が得られる。本発明では、この枚数がいくら大きくなっても、単に補間計算のステップ幅を変更すれば済むため、処理にまったく困難

な部分はない。現在、一秒間に3000枚のフレームを撮影できるほどのカメラは存在せず、本実施の形態によれば、任意の時間レートのスーパースローモーション画像が得られる。

【0175】

実施の形態5

本発明はデータ量を削減して動画を生成できるため、一般にCPUパワー、電池寿命、メモリ容量に制約がありがちなモバイル機器への応用にも有利である。そのため、中間画像の生成を、モバイル機器への適用を前提とした機能のもとで行ってもよい。

【0176】

その機能の例として、第1画像または第2画像の一方のみと前記対応情報をもとに生成された中間画像を表示してもよい。その場合、画像をテンポラリにしておくためのメモリの容量が少なくて済む。対応情報と第1画像のみで中間画像を生成する場合、第1画像の各画素やメッシュ上のポリゴンその他の画像要素（以下単に画素で代表させる）を対応情報にしたがって移動させればよい。すなわち、通常は対応情報で画素の位置および画素値すなわち色の両方を補間できるが、これを画素の位置の移動のみの処理に簡素化する。

【0177】

別の機能として、第1画像のみと対応情報をストリーミング処理して中間画像を生成してもよい。一例として、

1. $t = 0$ の画像 $I(0)$ と対応情報から $t = 0 \sim 1$ の中間画像を生成して表示し、画像 $I(0)$ を捨てる
2. $t = 1$ の画像 $I(1)$ と対応情報から $t = 1 \sim 2$ の中間画像を生成して表示し、画像 $I(1)$ を捨てる

という処理を繰り返すことにより、比較的小さなメモリ容量で動画のストリーミング表示が実現する。いずれの場合も、中間画像は表示の瞬間だけフレームバッファに存在すればよい。

【0178】

以上、本発明をいくつかの実施の形態をもとに説明した。これらの機能または

処理は、いずれもコンピュータプログラムの形で提供することができる。また、以上の処理の任意の組合せや組み替えもまた、本発明の態様として有効である。

【図面の簡単な説明】

【図 1】 図 1 (a) とは図 1 (b) は、ふたりの人物の顔に平均化フィルタを施して得られる画像、図 1 (c) と図 1 (d) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 0)$ の画像、図 1 (e) と図 1 (f) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 1)$ の画像、図 1 (g) と図 1 (h) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 2)$ の画像、図 1 (i) と図 1 (j) は、ふたりの人物の顔に関して前提技術で求められる $p(5, 3)$ の画像をそれぞれディスプレイ上に表示した中間調画像の写真である。

【図 2】 図 2 (R) はもとの四辺形を示す図、図 2 (A)、図 2 (B)、図 2 (C)、図 2 (D)、図 2 (E) はそれぞれ相続四辺形を示す図である。

【図 3】 始点画像と終点画像の関係、および第 m レベルと第 $m-1$ レベルの関係を相続四辺形を用いて示す図である。

【図 4】 パラメータ η とエネルギー C_f の関係を示す図である。

【図 5】 図 5 (a)、図 5 (b) は、ある点に関する写像が全単射条件を満たすか否かを外積計算から求める様子を示す図である。

【図 6】 前提技術の全体手順を示すフローチャートである。

【図 7】 図 6 の S 1 の詳細を示すフローチャートである。

【図 8】 図 7 の S 1 0 の詳細を示すフローチャートである。

【図 9】 第 m レベルの画像の一部と、第 $m-1$ レベルの画像の一部の対応関係を示す図である。

【図 1 0】 前提技術で生成された始点階層画像を示す図である。

【図 1 1】 図 6 の S 2 に進む前に、マッチング評価の準備の手順を示す図である。

【図 1 2】 図 6 の S 2 の詳細を示すフローチャートである。

【図 1 3】 第 0 レベルにおいて副写像を決定する様子を示す図である。

【図 1 4】 第 1 レベルにおいて副写像を決定する様子を示す図である。

【図 15】 図 12 の S 2 1 の詳細を示すフローチャートである。

【図 16】 ある $f(m, s)$ について λ を変えながら求められた $f(m, s)$ ($\lambda = i \Delta \lambda$) に対応するエネルギー $C(m, s)_f$ の挙動を示す図である。

【図 17】 η を変えながら求められた $f(n)$ ($\eta = i \Delta \eta$) ($i = 0, 1, \dots$) に対応するエネルギー $C(n)_f$ の挙動を示す図である。

【図 18】 第 1 画像と第 2 画像間である画素の対応する様子を示す図である。

【図 19】 第 1 画像上にとられた起点ポリゴンと第 2 画像上にとられた終点ポリゴンの対応関係を示す図である。

【図 20】 起点ポリゴン内の点に対応する終点ポリゴン内の点を求める手順を示す図である。

【図 21】 実施の形態 1 で対応点ファイルを生成する手順を示すフローチャートである。

【図 22】 実施の形態 1 で、対応点ファイルをもとに中間画像を生成する手順を示すフローチャートである。

【図 23】 実施の形態 1 に係る映像効果装置の構成図である。

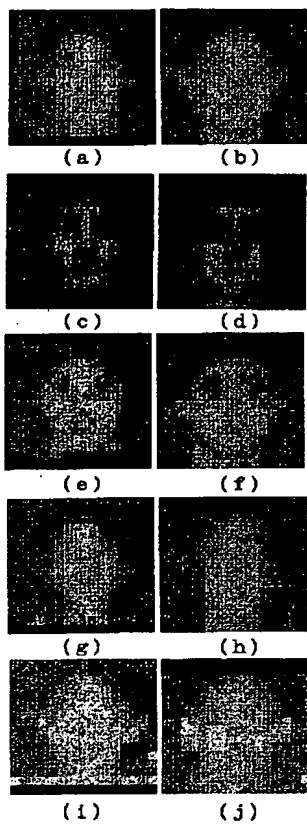
【符号の説明】

- 10 映像効果装置
- 12 画像データ入力部
- 14 マッチングプロセッサ
- 16 対応点ファイル保持部
- 18 中間画像生成部
- 20 表示部
- 22 通信部

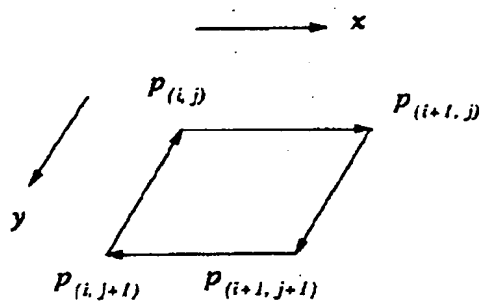
【書類名】

図面

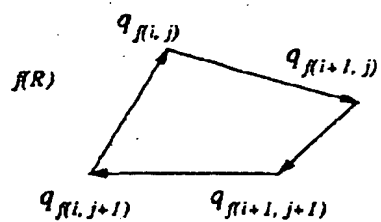
【図 1】



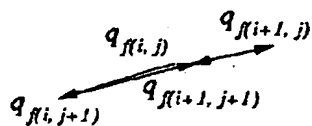
【図2】



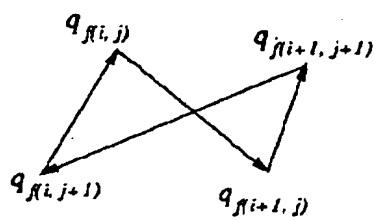
(R)



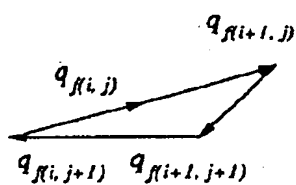
(A)



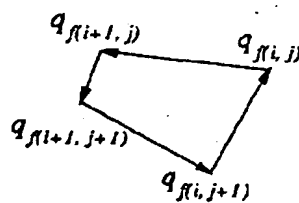
(E)



(B)

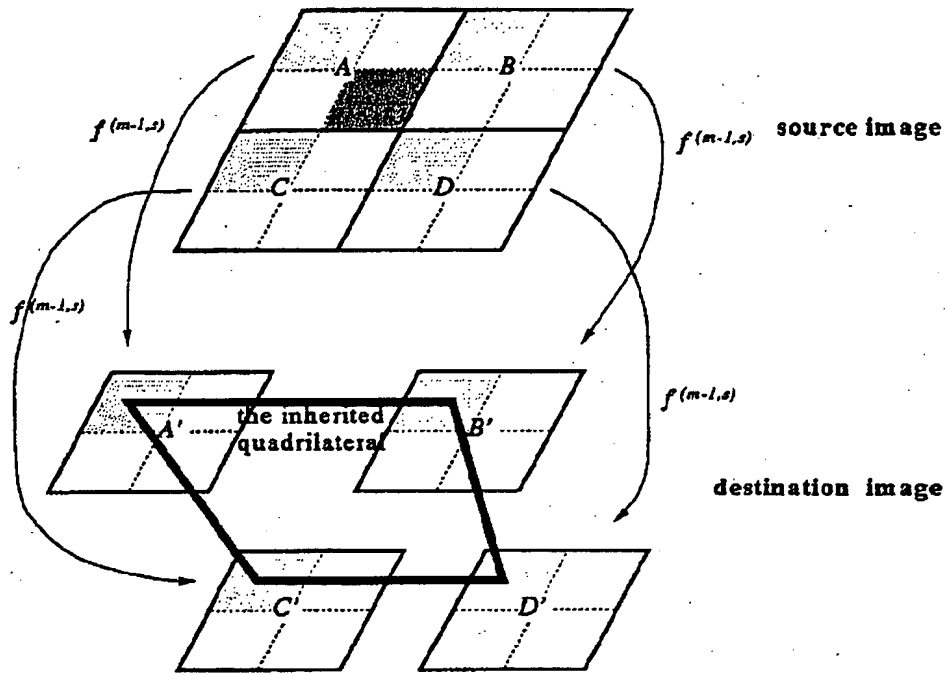


(D)

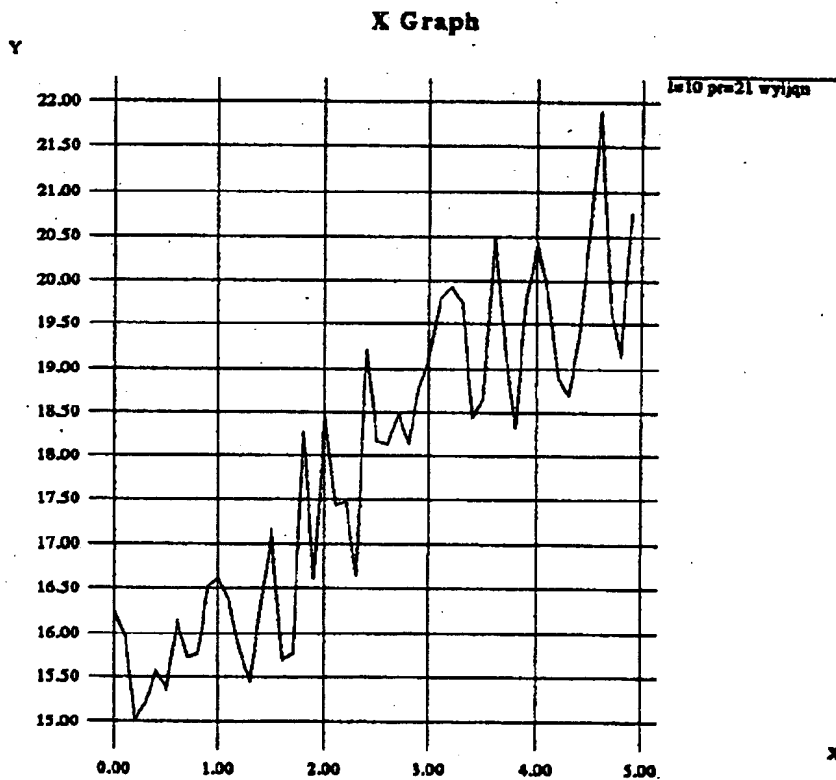


(C)

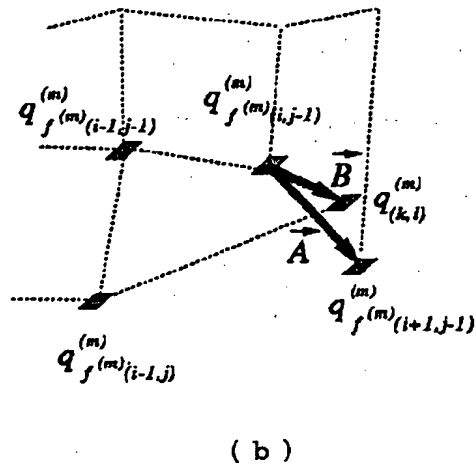
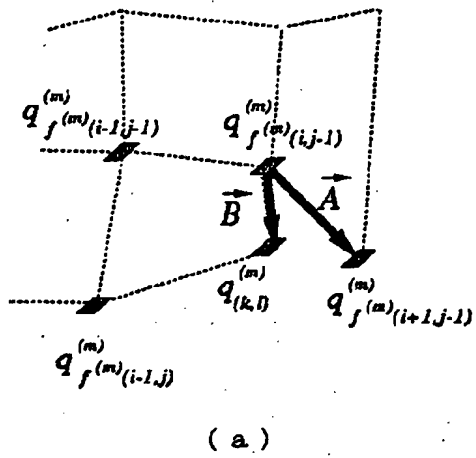
【図 3】



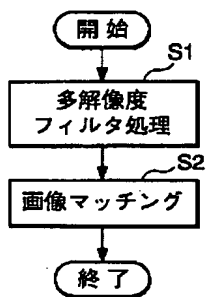
【図 4】



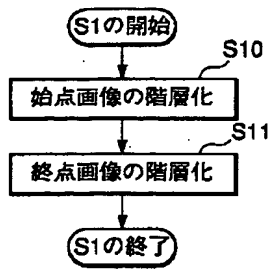
【図 5】



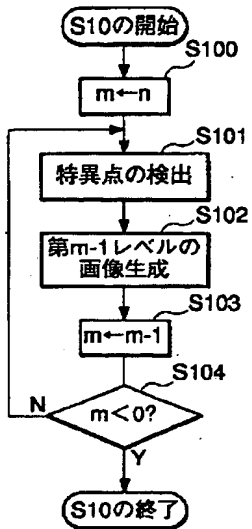
【図 6】



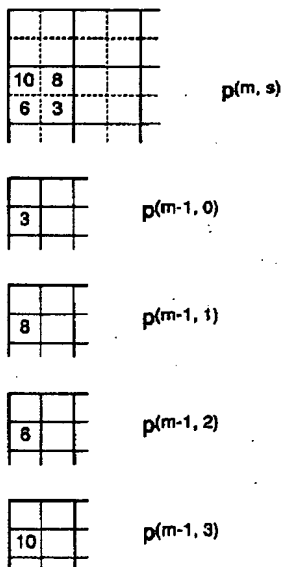
【図 7】



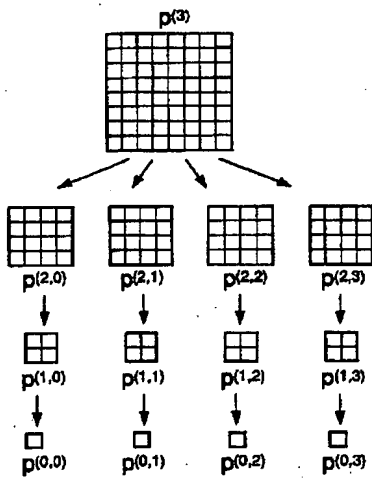
【図 8】



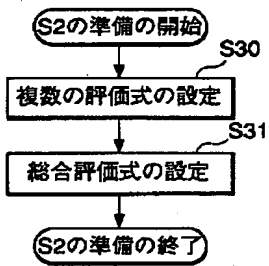
【図 9】



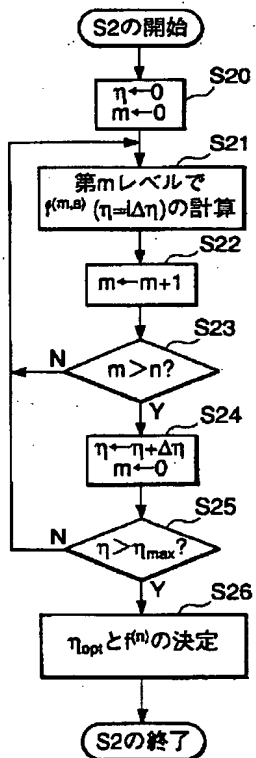
【図 10】



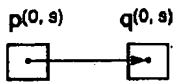
【図 11】



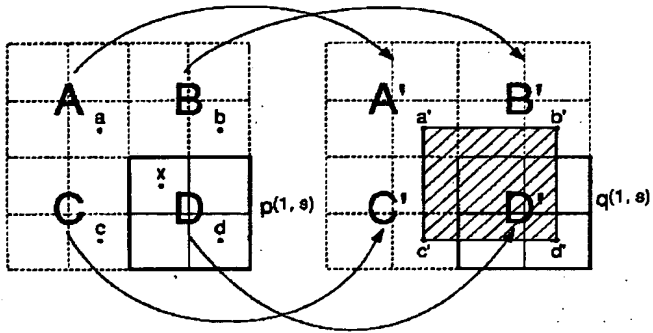
【図 12】



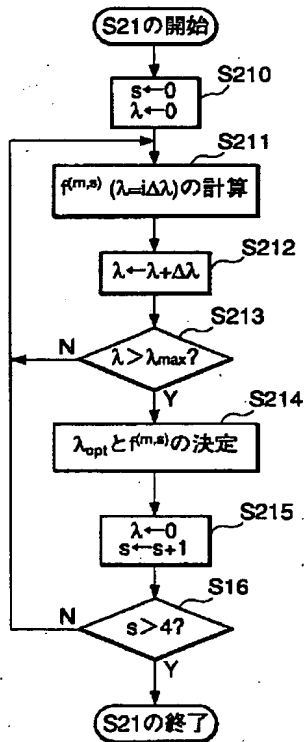
【図 13】



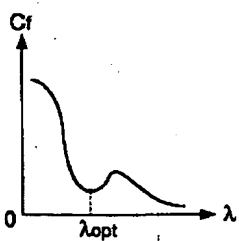
【図 14】



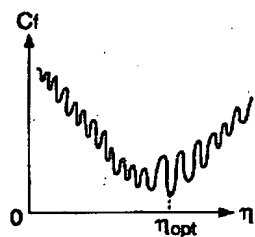
【図 15】



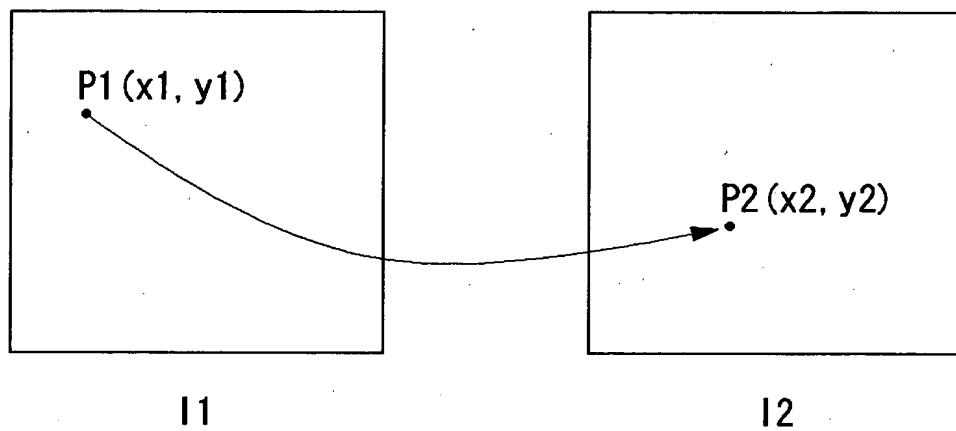
【図 16】



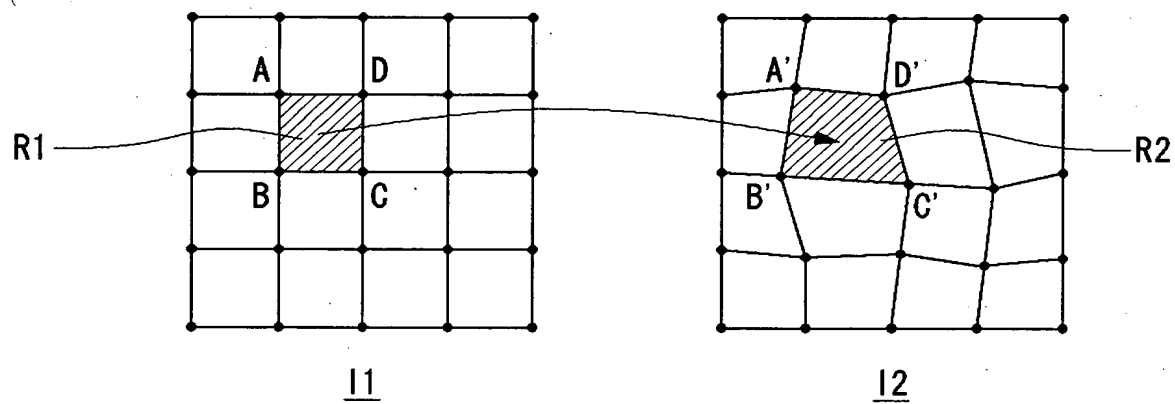
【図 17】



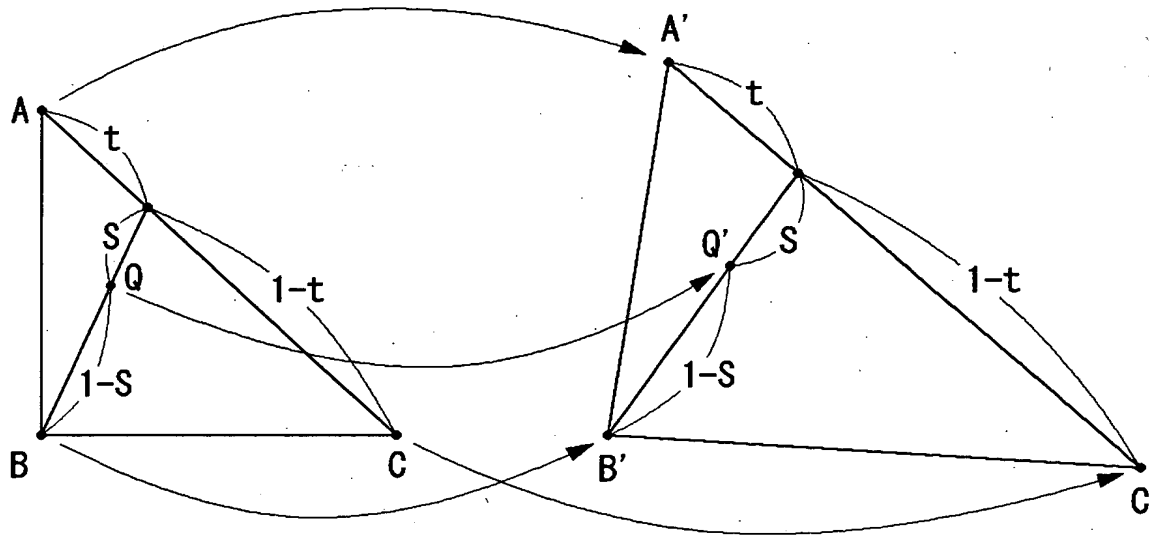
【図 18】



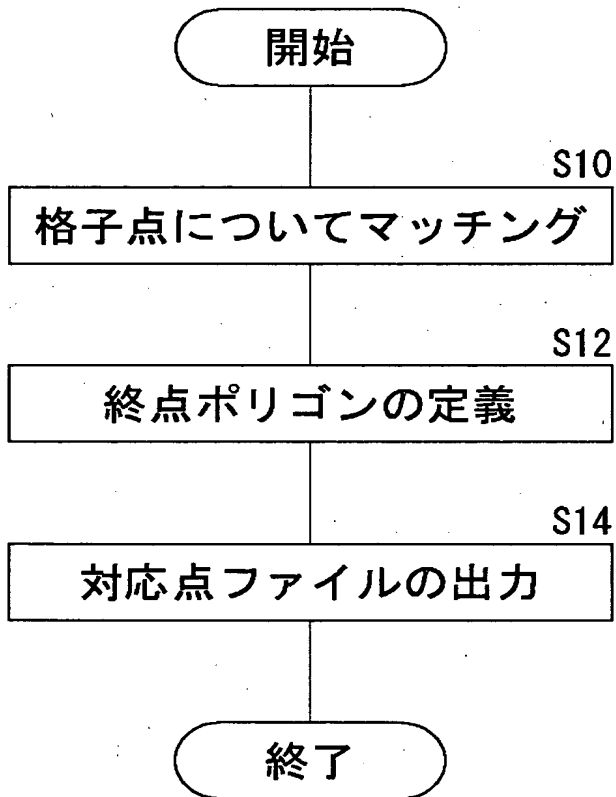
【図 19】



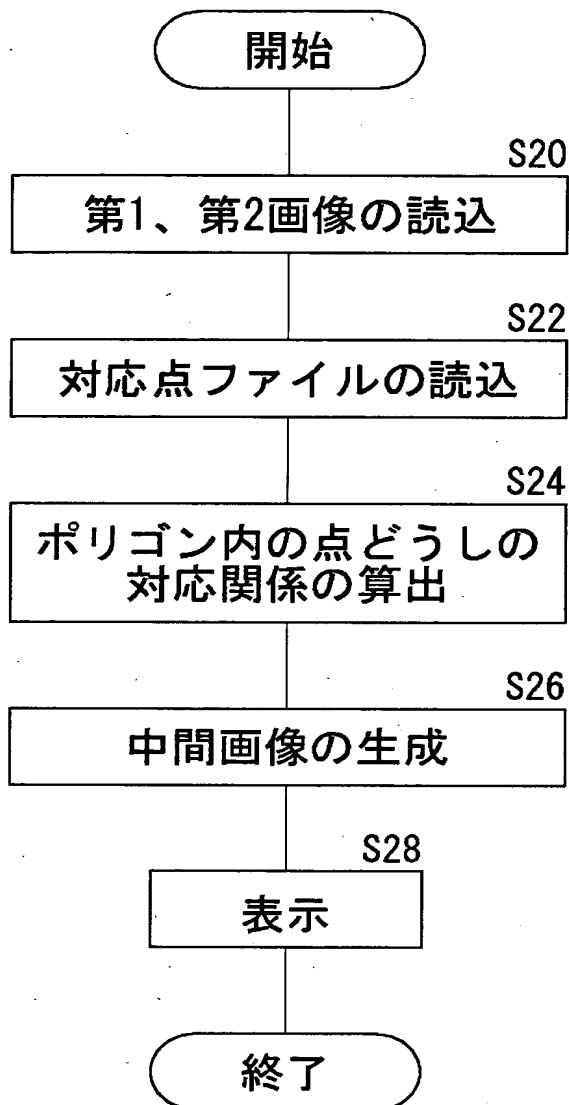
【図20】



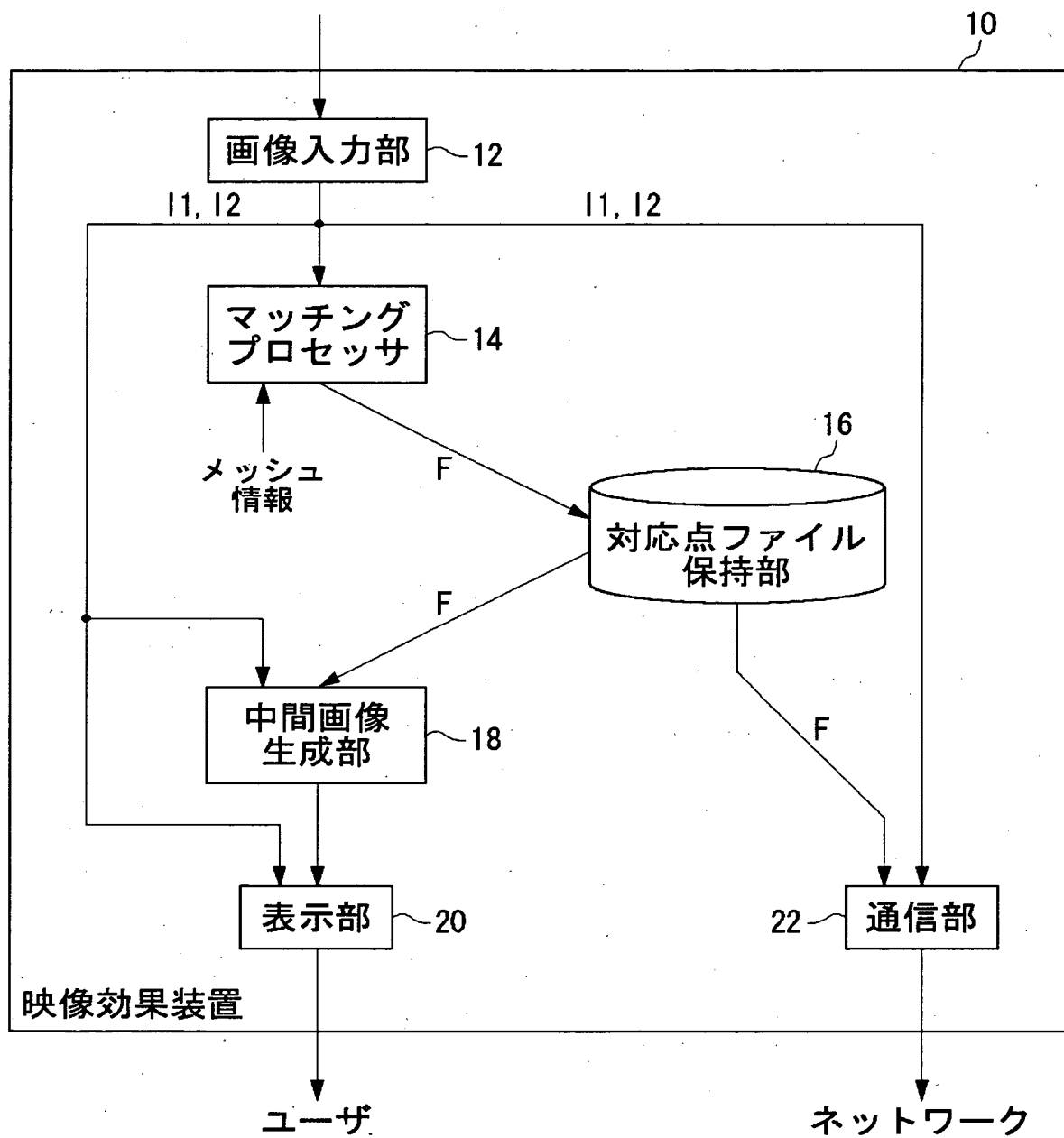
【図21】



【図 2 2】



【図 23】



【書類名】 要約書

【要約】

【課題】 映像効果を簡便に、高い品質で、しかも軽いデータで実現することが望まれつつ、困難な要請だった。

【解決手段】 画像入力部12は第1画像と第2画像を入力する。マッチングプロセッサ14はそれらの画像間で画素マッチングを行い、第1画像上に取られたメッシュの格子点について、その対応点を第2画像上に求める。その結果が対応点ファイルFとして記録される。中間画像生成部18は、対応点ファイルFをもとに第1画像と第2画像間の中間画像を生成する。格子点だけで対応点ファイルを作ったので、そのデータ容量は非常に小さくなり、簡便に各種映像効果、動画圧縮、オンライン商品紹介、携帯機器への実装が実現する。

【選択図】 図23

出 願 人 履 歴 情 報

識別番号 [396001980]

1. 変更年月日 1996年 1月26日

[変更理由] 新規登録

住 所 東京都港区麻布十番1丁目7番3号

氏 名 株式会社モノリス